167

GNE.3230R1C39 PATENT

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | |
|---|---|---|
| Applicant | : | Eaton, et al. |
| Appl. No. | : | 10/063,557 |
| Filed | : | May 2, 2002 |
| For | : | SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME |
| Examiner | : | David J. Blanchard |
| Group Art Unit | : | 1642 |

**DECLARATION OF J. CHRISTOPHER GRIMALDI, UNDER 37 CFR §1.132**

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

I, J. Christopher Grimaldi, declare and state as follows:

1. I am a Senior Research Associate in the Molecular Biology Department of Genentech, Inc., South San Francisco, CA 94080.

2. My scientific Curriculum Vitae, including my list of publications, is attached to and forms part of this Declaration (Exhibit A).

3. I joined Genentech in January of 1999. From 1999 to 2003, I directed the Cloning Laboratory in the Molecular Biology Department. During this time I directed or performed numerous molecular biology techniques including semi-quantitative Polymerase Chain Reaction (PCR) analyses. I am currently involved, among other projects, in the isolation of genes coding for membrane associated proteins which can be used as targets for antibody therapeutics against cancer. In connection with the above-identified patent application, I personally performed or directed the semi-quantitative PCR gene expression analyses in the assay entitled "Tumor Versus Normal Differential Tissue Expression Distribution," which is described in EXAMPLE 18 in the specification. These studies were used to identify differences in gene expression between tumor tissue and their normal counterparts.

4. EXAMPLE 18 reports the results of the PCR analyses conducted as part of the investigating of several newly discovered DNA sequences. This process included developing

primers and analyzing expression of the DNA sequences of interest in normal and tumor tissues. The analyses were designed to determine whether a difference exists between gene expression in normal tissues as compared to tumor in the same tissue type.

5. The DNA libraries used in the gene expression studies were made from pooled samples of normal and of tumor tissues. Data from pooled samples is more likely to be accurate than data obtained from a sample from a single individual. That is, the detection of variations in gene expression is likely to represent a more generally relevant condition when pooled samples from normal tissues are compared with pooled samples from tumors in the same tissue type.

6. In differential gene expression studies, one looks for genes whose expression levels differ significantly under different conditions, for example, in normal versus diseased tissue. Thus, I conducted a semi-quantitative analysis of the expression of the DNA sequences of interest in normal versus tumor tissues. Expression levels were graded according to a scale of +, -, and +/- to indicate the amount of the specific signal detected. Using the widely accepted technique of PCR, it was determined whether the polynucleotides tested were more highly expressed, less expressed, or whether expression remained the same in tumor tissue as compared to its normal counterpart. Because this technique relies on the visual detection of ethidium bromide staining of PCR products on agarose gels, it is reasonable to assume that any detectable differences seen between two samples will represent at least a two fold difference in cDNA.

7. The results of the gene expression studies indicate that the genes of interest can be used to differentiate tumor from normal. The precise levels of gene expression are irrelevant; what matters is that there is a relative difference in expression between normal tissue and tumor tissue. The precise type of tumor is also irrelevant; again, the assay was designed to indicate whether a difference exists between normal tissue and tumor tissue of the same type. If a difference is detected, this indicates that the gene and its corresponding polypeptide and antibodies against the polypeptide are useful for diagnostic purposes, to screen samples to differentiate between normal and tumor. Additional studies can then be conducted if further information is desired.

8. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful statements may jeopardize the validity of the application or any patent issued thereon.

By: _____ Date: 8/10/2004

J. Christopher Grimaldi

S:\DOCS\AOK\AOK-5591.DOC/081004

# EXHIBIT A

# J. Christopher Grimaldi

1434-36th Ave.
San Francisco, CA 94122
(415) 681-1639 (Home)


**EDUCATION** University of California, Berkeley
Bachelor of Arts in Molecular Biology, 1984

## EMPLOYMENT EXPERIENCE

SRA Genentech Inc., South San Francisco; 1/99 to present

Previously, was responsible to direct and manage the Cloning Lab. Currently focused on isolating cancer specific genes for the Tumor Antigen (TAP), and Secreted Tumor Protein (STOP) projects for the Oncology Department as well as Immunologically relevant genes for the Immunology Department. Directed a lab of 6 scientists focused on a company-wide team effort to identify and isolate secreted proteins for potential therapeutic use (SPDI). For the SPDI project my duties were, among other things, the critically important coordination of the cloning of thousands of putative genes, by developing a smooth process of communication between the Bioinformatics, Cloning, Sequencing, and Legal teams. Collaborated with several groups to discover novel genes through the Curagen project, a unique differential display methodology. Interacted extensively with the Legal team providing essential data needed for filing patents on novel genes discovered through the SPDI, TAP and Curagen projects. My group has developed, implemented and patented high throughput cloning methodologies that have proven to be essential for the isolation of hundreds of novel genes for the SPDI, TAP and Curagen projects as well as dozens of other smaller projects.

Scientist DNAX Research Institute, Palo Alto; 9/91 to 1/99

Involved in multiple ~~projects~~ aimed at understanding novel genes discovered through bioinformatics studies and functional assays. Developed and patented a method for the specific depletion of eosinophils in vivo using monoclonal antibodies. Developed and implemented essential technical methodologies and provided strategic direction in the areas of expression, cloning, protein purification, general molecular biology, and monoclonal antibody production. Trained and supervised numerous technical staff.

Facilities Manager Corixa, Redwood City; 5/89 - 7/91.

Directed plant-related activities, which included expansion planning, maintenance, safety, purchasing, inventory control, shipping and receiving, and laboratory management. Designed and implemented the safety program. Also served as liaison to regulatory agencies at the local, state and federal level. Was in charge of property leases, leasehold improvements, etc. Negotiated vendor contracts and directed the purchasing department. Trained and supervised personnel to carry out the above-mentioned duties.

SRA University of California, San Francisco
Cancer Research Institute; 2/87-4/89.

Was responsible for numerous cloning projects including: studies of somatic hypermutation, studies of AIDS-associated lymphomas, and cloning of t(5;14), t(11;14), and t(8;14) translocations. Focused on the activation of hemopoietic growth factors involved in the t(5;14) translocation in leukemia patients..

Research
Technician Berlex Biosciences, South San Francisco; 7/85-2/87.

Worked on a subunit porcine vaccine directed against Mycoplasma hyopneumoniae. Was responsible for generating genomic libraries, screening with degenerate oligonucleotides, and characterizing and expressing clones in E. coli. Also constructed a general purpose expression vector for use by other scientific teams.

## PUBLICATIONS


1. Hilary F. Clark, et al. "The Secreted Protein Discovery Initiative (SPDI), a Large-scale Effort to Identify Novel Human Secreted and Transmembrane Proteins: a bioinformatics assessment." Genome Res. Vol 13(10), 2265-2270, 2003

2. Sean H. Adams, Clarissa Chui Sarah L. Schilbach, Xing Xian Yu, Audrey D. Goddard, J. Christopher Grimaldi, James Lee, Patrick Dowd, David A. Lewin, & Steven Colman "BFIT, a Unique Acyl-CoA Thioesterase Induced in Thermogenic Brown Adiopose Tissue: Cloning, organization of the humanb gene and assessment of a potential link to obesity" Biochemical Journal, Vol 360, 135-142, 2001

3. Szeto W, Jiang W, Tice DA, Rubinfeld B, Hollingshead PG, Fong SE, Dugger DL, Pham T, Yansura D, Wong TA, Grimaldi JC, Corpuz RT, Singh JS, Frantz GD, , Devaux B, Crowley CW, Schwall RH, Eberhard DA, Rastelli L, Polakis P, and Rennica D. "Overexpression of the Retenoic Acid-Responsive Gene Stra6 in Human Cancers and its Synergistic Activation by Wnt-1 and Retinoic Acid." Cancer Research Vol. 61(10), 4197-4205, 2001

4. Jeanne Kahn, Fuad Mehraban, Gladdys Ingle, Xiaohua Xin, Juliet E. Bryant, Gordon Vehar, Jill Schoenfeld, J. Christopher Grimaldi (incorrectly named as "Grimaldi, CJ"), Franklin Peale, Aparna Draksharapu, David A. Lewin, and Mary E. Gerritsen. "Gene Expression Profiling in an in Vitro Model of Angiogenesis." American Journal of Pathology Vol 156(6), 1887-1900, 2000.

5. Grimaldi JC, Yu NX, Grunig G, Seymour BW, Cottrez F, Robinson DS, Hosken N, Ferlin WG, Wu X, Soto H, O'Garra A, Howard MC, Coffman RL. "Depletion of eosinophils in mice through the use of antibodies specific for C-C chemokine receptor 3 (CCR3). Journal of Leukocyte Biology; Vol. 65(6), 846-53, 1999

6. Oliver AM, Grimaldi JC, Howard MC, Kearney JF. "Independently ligating CD38 and Fc gammaRIIB relays a dominant negative signal to B cells." Hybridoma Vol. 18(2), 113-9, 1999

7. Cockayne DA, Muchamuel T, Grimaldi JC, Muller-Steffner H, Randall TD, Lund FE, Murray R, Schuber F, Howard MC. "Mice deficient for the ecto-nicotinamide adenine dinucleotide glycohydrolase CD38 exhibit altered humoral immune responses." Blood Vol. 92(4), 1324-33, 1998

8. Frances E. Lund, Nanette W. Solvason, Michael P. Cooke, Andrew W. Heath, J. Christopher Grimaldi, Troy D. Randall, R. M. E. Parkhouse, Christopher C Goodnow and Maureen C. Howard. "Signaling through murine CD38 is impaired in antigen receptor unresponsive B cells." European Journal of Immunology, Vol. 25(5), 1338-1345, 1995

9. M. J. Guimaraes, J. F. Bazan, A. Zolotnik, M. V. Wiles, J. C. Grimaldi, F. Lee, T. McClanahan. "A new approach to the study of haematopoietic development in the yolk sac and embryoid body." Development, Vol. 121(10), 3335-3346, 1995

10. J. Christopher Grimaldi, Sriram Balasubramanian, J. Fernando Bazan, Armen Shanafelt, Gerard Zurawski and Maureen Howard. "CD38-mediated protein ribosylation." Journal of Immunology,Vol. 155(2), 811-817, 1995

11. Leopoldo Santos-Argumedo, Frances F. Lund, Andrew W. Heath, Nanette Solvason, Wei Wei Wu, J. Christopher Grimaldi, R. M. E. Parkhouse and Maureen Howard. "CD38 unresponsiveness of xid B cells implicates Bruton's tyrosine kinase (btk) as a regulator of CD38 induced signal transduction." International Immunology, Vol 7(2), 163-170, 1995

12. Frances Lund, Nanette Solvason, J. Christopher Grimaldi, R. M. E. Parkhouse and Maureen Howard. "Murine CD38: An immunoregulatory ectoenzyme." Immunology Today, Vol. 16(10), 469-473, 1995

13. Maureen Howard, J. Christopher Grimaldi, J. Fernando Bazan, Frances E. Lund, Leopoldo Santos-Argumedo, R. M. E. Parkhouse, Timothy F. Walseth, and Hon Cheung Lee. "Formation and Hydrolysis of Cyclic ADP-Ribose Catalyzed by Lymphocyte Antigen CD38." Science, Vol. 262, 1056-1059, 1993

14. Nobuyuki Harada, Leopoldo Santos-Argumedo, Ray Chang, J. Christopher Grimaldi, Frances Lund, Camilynn I. Brannan, Neal G. Copeland, Nancy A. Jenkins, Andrew Heath, R. M. E. Parkhouse and Maureen Howard. "Expression Cloning of a cDNA Encoding a Novel Murine B Cell Activation Marker: Homology to Human CD38." The Journal of Immunology, Vol. 151, 3111-3118, 1993

15. David J. Rawlings, Douglas C. Saffran, Satoshi Tsukada, David A. Largaespada, J. Christopher Grimaldi, Lucie Cohen Randolph N. Mohr, J. Fernando Bazan, Maureen Howard, Neal G. Copeland, Nancy A. Jenkins, Owen Witte. "Mutation of Unique Region of Bruton's Tyrosine Kinase in Immunodeficient XID Mice." Science, Vol. 261, 358-360, 1993

16. J. Christopher Grimaldi, Raul Torres, Christine A. Kozak, Ray Chang, Edward Clark, Maureen Howard, and Debra A. Cockayne. "Genomic Structure and Chromosomal Mapping of the Murine CD40 Gene." The Journal of Immunology, Vol 149, 3921-3926, 1992

17. Timothy C. Meeker, Bruce Shiramizu, Lawrence Kaplan, Brian Herndier, Henry Sanchez, J. Christopher Grimaldi, James Baumgartner, Jacab Rachlin, Ellen Feigal, Mark Rosenblum and Michael S. McGrath. "Evidence for Molecular Subtypes of HIV-Associated Lymphoma:

Division into Peripheral Monoclonal, Polyclonal and Central Nervous System Lymphoma." AIDS, Vol. 5, 669-674, 1991

18. Ann Grimaldi and Chris Grimaldi. "Small-Scale Lambda DNA Prep." Contribution to Current Protocols in Molecular Biology, Supplement 5, Winter 1989

19. J. Christopher Grimaldi, Timothy C. Meeker. "The t(5;14) Chromosomal Translocation in a Case of Acute Lymphocytic Leukemia Joins the Interleukin-3 Gene to the Immunoglobulin Heavy Chain Gene." Blood, Vol. 73, 2081-2085, 1989

20. Timothy C. Meeker, J. Christopher Grimaldi, et al. "An Additional Breakpoint Region in the BCL-1 Locus Associated with the t(11;14) (q13;q32) Translocation of B-Lymphocytic Malignancy." Blood, Vol. 74, 1801-1806, 1989

21 Timothy C. Meeker, J. Christopher Grimaldi, Robert O'Rourke, et al. "Lack of Detectable Somatic Hypermutation in the V Region of the Ig H Chain Gene of a Human Chronic B Lymphocytic Leukemia." The Journal of Immunology, Vol. 141, 3994-3998, 1988

## MANUSCRIPTS IN PREPARATION

1. Sriram Balasubramanian, J. Christopher Grimaldi, J. Fernando Bazan, Gerard Zurawski and Maureen Howard. "Structural and functional characterization of CD38: Identification of active site residues"

## PATENTS

1. "Methods for Eosinophil Depletion with Antibody to CCR3 Receptor" (US 6,207,155 B1).

2. "Amplification Based Cloning Method." (US 6,607,899)

3. Ashkenazi et al., "Secreted and Transmembrane Polypeptides and Nucleic Acids Encoding the Same." (this patent covers several hundred genes)

4. "IL-17 Homologous Polypeptides and Therapeutic Uses Thereof"

5. "Method of Diagnosing and Treating Cartilaginous Disorders."

## MEMBERSHIPS AND ACTIVITIES

| | |
|---|---|
| Editor | Frontiers in Bioscience |
| Member | DNAX Safety Committee 1991-1999<br>Biological Safety Affairs Forum (BSAF) 1990-1991<br>Environmental Law Foundation (ELF) 1990-1991 |

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | |
|---|---|---|
| Applicant | : | Eaton, et al. |
| Appl. No. | : | 10/063,557 |
| Filed | : | May 2, 2002 |
| For | : | SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME |
| Examiner | : | David J. Blanchard |
| Group Art Unit | : | 1642 |

**DECLARATION OF J. CHRISTOPHER GRIMALDI, UNDER 37 C.F.R. § 1.132**

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

I, J. Christopher Grimaldi, declare and say as follows:

1. I am a Senior Research Associate in the Molecular Biology Department of Genentech, Inc., South San Francisco, CA 94080.

2. I joined Genentech in January of 1999. From 1999 to 2003, I directed the Cloning Laboratory in the Molecular Biology Department. During this time I directed or performed numerous molecular biology techniques including qualitative Polymerase Chain Reaction (PCR) analyses. I am currently involved in, among other projects, the isolation of genes coding for membrane associated proteins which can be used as targets for antibody therapeutics against cancer. In connection with the above-identified patent application, I personally performed or directed the semi-quantitative PCR analyses in the assay entitled "Tumor Versus Normal Differential Tissue Expression Distribution" which is described in EXAMPLE 18 in the specification that were used to identify differences in gene expression between tumor tissue and their normal counterparts.

3. My scientific Curriculum Vitae, including my list of publications, is attached to and forms part of this Declaration (Exhibit A).

4. In differential gene expression studies, one looks for genes whose expression levels differ significantly under different conditions, for example, in normal versus diseased tissue.

Chromosomal aberrations, such as gene amplification, and chromosomal translocations are important markers of specific types of cancer and lead to the aberrant expression of specific genes and their encoded polypeptides, including over-expression and under-expression. For example, gene amplification is a process in which specific regions of a chromosome are duplicated, thus creating multiple copies of certain genes that normally exist as a single copy. Gene under-expression can occur when a gene is not transcribed into mRNA. In addition, chromosomal translocations occur when two different chromosomes break and are rejoined to each other chromosome resulting in a chimeric chromosome which displays a different expression pattern relative to the parent chromosomes. Amplification of certain genes such as Her2/Neu [Singleton *et al.*, Pathol. Annu., 27Pt1:165-190], or chromosomal translocations such as t(5;14), [Grimaldi *et al.*, Blood, 73(8):2081-2085(1989); Meeker *et al.*, Blood, 76(2):285-289(1990)] give cancer cells a growth or survival advantage relative to normal cells, and might also provide a mechanism of tumor cell resistance to chemotherapy or radiotherapy. When the chromosomal aberration results in the aberrant expression of a mRNA and the corresponding gene product (the polypeptide), as it does in the aforementioned cases, the gene product is a promising target for cancer therapy, for example, by the therapeutic antibody approach.

5. Comparison of gene expression levels in normal versus diseased tissue has important implications both diagnostically and therapeutically. For example, those who work in this field are well aware that in the vast majority of cases, when a gene is over-expressed, as evidenced by an increased production of mRNA, the gene product or polypeptide will also be over-expressed. It is unlikely that one identifies increased mRNA expression without associated increased protein expression. This same principle applies to gene under-expression. When a gene is under-expressed, the gene product is also likely to be under-expressed. Stated in another way, two cell samples which have differing mRNA concentrations for a specific gene are expected to have correspondingly different concentration of protein for that gene. Techniques used to detect mRNA, such as Northern Blotting, Differential Display, *in situ* hybridization, quantitative PCR, Taqman, and more recently Microarray technology all rely on the dogma that a change in mRNA will represent a similar change in protein. If this dogma did not hold true then these techniques would have little value and not be so widely used. The use of mRNA quantitation techniques have identified a seemingly endless number of genes which are differentially expressed in various tissues and these genes have subsequently been shown to have correspondingly similar changes in their protein levels. Thus, the detection of increased mRNA expression is expected to result in increased polypeptide expression, and the detection of decreased mRNA expression is expected to result in decreased polypeptide expression. The detection of increased or decreased polypeptide expression can be used for cancer diagnosis and treatment.

6. However, even in the rare case where the protein expression does not correlate with the mRNA expression, this still provides significant information useful for cancer diagnosis and treatment. For example, if over- or under-expression of a gene product does not correlate with over- or under-expression of mRNA in certain tumor types but does so in others, then identification of both gene expression and protein expression enables more accurate tumor classification and hence better determination of suitable therapy. In addition, absence of over- or

under-expression of the gene product in the presence of a particular over- or under-expression of mRNA is crucial information for the practicing clinician. For example, if a gene is over-expressed but the corresponding gene product is not significantly over-expressed, the clinician accordingly will decide not to treat a patient with agents that target that gene product.

7. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful statements may jeopardize the validity of the application or any patent issued thereon.

By: _____ Date: 8/10/2004

J. Christopher Grimaldi

S:\DOCS\AOK\AOK-5479.DOC
071904

# EXHIBIT A

# J. Christopher Grimaldi

1434-36th Ave.
San Francisco, CA 94122
(415) 681-1639 (Home)


**EDUCATION** University of California, Berkeley
Bachelor of Arts in Molecular Biology, 1984

## EMPLOYMENT EXPERIENCE

SRA — Genentech Inc., South San Francisco; 1/99 to present

Previously, was responsible to direct and manage the Cloning Lab. Currently focused on isolating cancer specific genes for the Tumor Antigen (TAP), and Secreted Tumor Protein (STOP) projects for the Oncology Department as well as Immunologically relevant genes for the Immunology Department. Directed a lab of 6 scientists focused on a company-wide team effort to identify and isolate secreted proteins for potential therapeutic use (SPDI). For the SPDI project my duties were, among other things, the critically important coordination of the cloning of thousands of putative genes, by developing a smooth process of communication between the Bioinformatics, Cloning, Sequencing, and Legal teams. Collaborated with several groups to discover novel genes through the Curagen project, a unique differential display methodology. Interacted extensively with the Legal team providing essential data needed for filing patents on novel genes discovered through the SPDI, TAP and Curagen projects. My group has developed, implemented and patented high throughput cloning methodologies that have proven to be essential for the isolation of hundreds of novel genes for the SPDI, TAP and Curagen projects as well as dozens of other smaller projects.

Scientist — DNAX Research Institute, Palo Alto; 9/91 to 1/99

Involved in multiple projects aimed at understanding novel genes discovered through bioinformatics studies and functional assays. Developed and patented a method for the specific depletion of eosinophils in vivo using monoclonal antibodies. Developed and implemented essential technical methodologies and provided strategic direction in the areas of expression, cloning, protein purification, general molecular biology, and monoclonal antibody production. Trained and supervised numerous technical staff.

Facilities Manager — Corixa, Redwood City; 5/89 - 7/91.

Directed plant-related activities, which included expansion planning, maintenance, safety, purchasing, inventory control, shipping and receiving, and laboratory management. Designed and implemented the safety program. Also served as liaison to regulatory agencies at the local, state and federal level. Was in charge of property leases, leasehold improvements, etc. Negotiated vendor contracts and directed the purchasing department. Trained and supervised personnel to carry out the above-mentioned duties.

SRA University of California, San Francisco
Cancer Research Institute; 2/87-4/89.

Was responsible for numerous cloning projects including: studies of somatic hypermutation, studies of AIDS-associated lymphomas, and cloning of t(5;14), t(11;14), and t(8;14) translocations. Focused on the activation of hemopoietic growth factors involved in the t(5;14) translocation in leukemia patients..

Research
Technician Berlex Biosciences, South San Francisco; 7/85-2/87.

Worked on a subunit porcine vaccine directed against Mycoplasma hyopneumoniae. Was responsible for generating genomic libraries, screening with degenerate oligonucleotides, and characterizing and expressing clones in E. coli. Also constructed a general purpose expression vector for use by other scientific teams.

**PUBLICATIONS**


1. Hilary F. Clark, et al. "The Secreted Protein Discovery Initiative (SPDI), a Large-scale Effort to Identify Novel Human Secreted and Transmembrane Proteins: a bioinformatics assessment." Genome Res. Vol 13(10), 2265-2270, 2003

2. Sean H. Adams, Clarissa Chui' Sarah L. Schilbach, Xing Xian Yu, Audrey D. Goddard, J. Christopher Grimaldi, James Lee, Patrick Dowd, David A. Lewin, & Steven Colman' "BFIT, a Unique Acyl-CoA Thioesterase Induced in Thermogenic Brown Adiopose Tissue: Cloning, organization of the humanb gene and assessment of a potential link to obesity" Biochemical Journal, Vol 360, 135-142, 2001

3. Szeto W, Jiang W, Tice DA, Rubinfeld B, Hollingshead PG, Fong SE, Dugger DL, Pham T, Yansura D, Wong TA, Grimaldi JC, Corpuz RT, Singh JS, Frantz GD, , Devaux B, Crowley CW, Schwall RH, Eberhard DA, Rastelli L, Polakis P, and Pennica D. "Overexpression of the Retenoic Acid-Responsive Gene Stra6 in Human Cancers and its Synergistic Activation by Wnt-1 and Retinoic Acid." Cancer Research Vol. 61(10), 4197-4205, 2001

4. Jeanne Kahn, Fuad Mehraban, Gladdys Ingle, Xiaohua Xin, Juliet E. Bryant, Gordon Vehar, Jill Schoenfeld, J. Christopher Grimaldi (incorrectly named as "Grimaldi, CJ"), Franklin Peale, Aparna Draksharapu, David A. Lewin, and Mary E. Gerritsen. "Gene Expression Profiling in an in Vitro Model of Angiogenesis." American Journal of Pathology Vol 156(6), 1887-1900, 2000.

5. Grimaldi JC, Yu NX, Grunig G, Seymour BW, Cottrez F, Robinson DS, Hosken N, Ferlin WG, Wu X, Soto H, O'Garra A, Howard MC, Coffman RL. "Depletion of eosinophils in mice through the use of antibodies specific for C-C chemokine receptor 3 (CCR3). Journal of Leukocyte Biology; Vol. 65(6), 846-53, 1999

6. Oliver AM, Grimaldi JC, Howard MC, Kearney JF. "Independently ligating CD38 and Fc gammaRIIB relays a dominant negative signal to B cells." Hybridoma Vol. 18(2), 113-9, 1999

7. Cockayne DA, Muchamuel T, Grimaldi JC, Muller-Steffner H, Randall TD, Lund FE, Murray R, Schuber F, Howard MC. "Mice deficient for the ecto-nicotinamide adenine dinucleotide glycohydrolase CD38 exhibit altered humoral immune responses." Blood Vol. 92(4), 1324-33, 1998

8. Frances E. Lund, Nanette W. Solvason, Michael P. Cooke, Andrew W. Heath, J. Christopher Grimaldi, Troy D. Randall, R. M. E. Parkhouse, Christopher C Goodnow and Maureen C. Howard. "Signaling through murine CD38 is impaired in antigen receptor unresponsive B cells." European Journal of Immunology, Vol. 25(5), 1338-1345, 1995

9. M. J. Guimaraes, J. F. Bazan, A. Zolotnik, M. V. Wiles, J. C. Grimaldi, F. Lee, T. McClanahan. "A new approach to the study of haematopoietic development in the yolk sac and embryoid body." Development, Vol. 121(10), 3335-3346, 1995

10. J. Christopher Grimaldi, Sriram Balasubramanian, J. Fernando Bazan, Armen Shanafelt, Gerard Zurawski and Maureen Howard. "CD38-mediated protein ribosylation." Journal of Immunology,Vol. 155(2), 811-817, 1995

11. Leopoldo Santos-Argumedo, Frances F. Lund, Andrew W. Heath, Nanette Solvason, Wei Wei Wu, J. Christopher Grimaldi, R. M. E. Parkhouse and Maureen Howard. "CD38 unresponsiveness of xid B cells implicates Bruton's tyrosine kinase (btk) as a regulator of CD38 induced signal transduction." International Immunology, Vol 7(2), 163-170, 1995

12. Frances Lund, Nanette Solvason, J. Christopher Grimaldi, R. M. E. Parkhouse and Maureen Howard. "Murine CD38: An immunoregulatory ectoenzyme." Immunology Today, Vol. 16(10), 469-473, 1995

13. Maureen Howard, J. Christopher Grimaldi, J. Fernando Bazan, Frances E. Lund, Leopoldo Santos-Argumedo, R. M. E. Parkhouse, Timothy F. Walseth, and Hon Cheung Lee. "Formation and Hydrolysis of Cyclic ADP-Ribose Catalyzed by Lymphocyte Antigen CD38." Science, Vol. 262, 1056-1059, 1993

14. Nobuyuki Harada, Leopoldo Santos-Argumedo, Ray Chang, J. Christopher Grimaldi, Frances Lund, Camilynn I. Brannan, Neal G. Copeland, Nancy A. Jenkins, Andrew Heath, R. M. E. Parkhouse and Maureen Howard. "Expression Cloning of a cDNA Encoding a Novel Murine B Cell Activation Marker: Homology to Human CD38." The Journal of Immunology, Vol. 151, 3111-3118, 1993

15. David J. Rawlings, Douglas C. Saffran, Satoshi Tsukada, David A. Largaespada, J. Christopher Grimaldi, Lucie Cohen Randolph N. Mohr, J. Fernando Bazan, Maureen Howard, Neal G. Copeland, Nancy A. Jenkins, Owen Witte. "Mutation of Unique Region of Bruton's Tyrosine Kinase in Immunodeficient XID Mice." Science, Vol. 261, 358-360, 1993

16. J. Christopher Grimaldi, Raul Torres, Christine A. Kozak, Ray Chang, Edward Clark, Maureen Howard, and Debra A. Cockayne. "Genomic Structure and Chromosomal Mapping of the Murine CD40 Gene." The Journal of Immunology, Vol 149, 3921-3926, 1992

17. Timothy C. Meeker, Bruce Shiramizu, Lawrence Kaplan, Brian Herndier, Henry Sanchez, J. Christopher Grimaldi, James Baumgartner, Jacab Rachlin, Ellen Feigal, Mark Rosenblum and Michael S. McGrath. "Evidence for Molecular Subtypes of HIV-Associated Lymphoma:

Division into Peripheral Monoclonal, Polyclonal and Central Nervous System Lymphoma." AIDS, Vol. 5, 669-674, 1991

18. Ann Grimaldi and Chris Grimaldi. "Small-Scale Lambda DNA Prep." Contribution to Current Protocols in Molecular Biology, Supplement 5, Winter 1989

19. J. Christopher Grimaldi, Timothy C. Meeker. "The t(5;14) Chromosomal Translocation in a Case of Acute Lymphocytic Leukemia Joins the Interleukin-3 Gene to the Immunoglobulin Heavy Chain Gene." Blood, Vol. 73, 2081-2085, 1989

20. Timothy C. Meeker, J. Christopher Grimaldi, et al. "An Additional Breakpoint Region in the BCL-1 Locus Associated with the t(11;14) (q13;q32) Translocation of B-Lymphocytic Malignancy." Blood, Vol. 74, 1801-1806, 1989

21 Timothy C. Meeker, J. Christopher Grimaldi, Robert O'Rourke, et al. "Lack of Detectable Somatic Hypermutation in the V Region of the Ig H Chain Gene of a Human Chronic B Lymphocytic Leukemia." The Journal of Immunology, Vol. 141, 3994-3998, 1988

## MANUSCRIPTS IN PREPARATION

1. Sriram Balasubramanian, J. Christopher Grimaldi, J. Fernando Bazan, Gerard Zurawski and Maureen Howard. "Structural and functional characterization of CD38: Identification of active site residues"

## PATENTS

1. "Methods for Eosinophil Depletion with Antibody to CCR3 Receptor" (US 6,207,155 B1).

2. "Amplification Based Cloning Method." (US 6,607,899)

3. Ashkenazi et al., "Secreted and Transmembrane Polypeptides and Nucleic Acids Encoding the Same." (this patent covers several hundred genes)

4. "IL-17 Homologous Polypeptides and Therapeutic Uses Thereof"

5. "Method of Diagnosing and Treating Cartilaginous Disorders."

## MEMBERSHIPS AND ACTIVITIES

| | |
|---|---|
| Editor | Frontiers in Bioscience |
| Member | DNAX Safety Committee 1991-1999<br>Biological Safety Affairs Forum (BSAF) 1990-1991<br>Environmental Law Foundation (ELF) 1990-1991 |

# The t(5;14) Chromosomal Translocation in a Case of Acute Lymphocytic Leukemia Joins the Interleukin-3 Gene to the Immunoglobulin Heavy Chain Gene

By J. Christopher Grimaldi and Timothy C. Meeker

**Chromosomal translocations have proven to be important markers of the genetic abnormalities central to the pathogenesis of cancer. By cloning chromosomal breakpoints one can identify activated proto-oncogenes. We have studied a case of B-lineage acute lymphocytic leukemia (ALL) that was associated with peripheral blood eosinophilia. The chromosomal translocation t(5;14) (q31;q32) from this sample was cloned and studied at the molecular level. This translocation joined the immunoglobulin heavy chain joining (Jh) region to the promotor region of the interleukin-3 (IL-3) gene in opposite transcriptional orientations. The data suggest that activation of the IL-3 gene by the enhancer of the immunoglobulin heavy chain gene may play a central role in the pathogenesis of this leukemia and the associated eosinophilia.**



KARYOTYPIC STUDIES of leukemia and lymphoma have identified frequent nonrandom chromosomal translocations. Some of these translocations juxtapose the immunoglobulin heavy chain (IgH) gene with important protooncogenes, such as c-*myc* and *bcl*-2.[1,2] In this way, the IgH gene can activate proto-oncogenes, resulting in disordered gene expression and a step in the development of cancer. The investigation of additional nonrandom translocations into the IgH locus allows us to identify new genes promoting the generation of leukemia and lymphoma.

A distinct subtype of acute lymphocytic leukemia (ALL) has been characterized by B-lineage phenotype, associated eosinophilia in the peripheral blood, and a t(5;14)(q31;q32) chromosomal translocation.[3,4] This syndrome probably occurs in <1% of all patients with ALL. We hypothesized that the cloning of the translocation characteristic of this leukemia might allow the identification of an important gene on chromosome 5 that plays a role in the evolution of this disease. In this report we demonstrate that the interleukin-3 gene (IL-3) and the IgH gene are joined by this translocation.

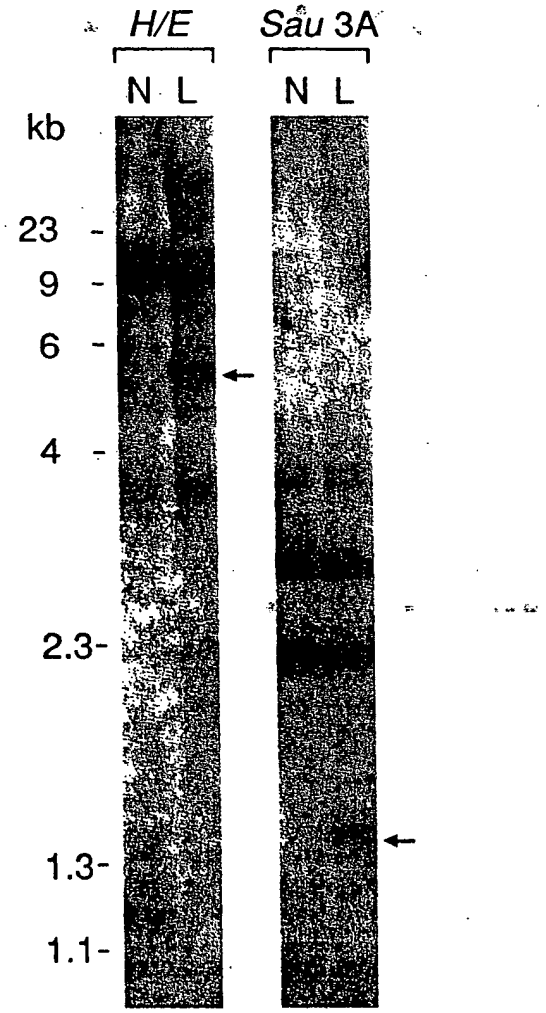



**Fig 1. DNA blots of the leukemia sample. The restriction fragment pattern of normal human DNA (N) and the leukemia sample (L) were compared using a human Jh probe. Rearranged bands are indicated by arrows. Sample L exhibits a single rearranged band with both *Hind* III/*Eco*RI and *Sau*3A restriction digests. The rearranged bands are less intense than the other bands because the majority of cells in the sample represent normal bone marrow elements.**

## MATERIALS AND METHODS

*Sample and DNA blots.* A bone marrow aspirate from a representative patient with ALL (L1 morphology by French-American-British [FAB] criteria), peripheral eosinophilia (up to 20,000 per microliter with a normal value of <350 per microliter) and a t(5;14)(q31;q32) translocation was studied. Using published methods, genomic DNA was isolated and DNA blots were made.[5] Briefly, 10 μg of high molecular weight (mol wt) DNA were digested using an appropriate restriction enzyme and electrophoresed on a 0.8% agarose gel. The gel was stained with ethidium bromide, photographed, denatured, neutralized, and transferred to Hybond (Amersham, Arlington Heights, IL). After treatment of the filter with ultraviolet light, hybridization was performed. The filter was washed to a final stringency of 0.2% saturated sodium citrate (SSC) and 0.1% sodium lauryl sulfate (SDS) and exposed to film. The human Jh probe has been previously reported.[6]

*Genomic library.* The genomic library was made using pub-

*From the Division of Hematology/Oncology, Department of Medicine, University of California, San Francisco.*

*Submitted February 22, 1989; accepted March 8, 1989.*

*Supported by NIH Grant No. CA01102.*

*Address reprint requests to Timothy C. Meeker, MD, UCSF/VAMC 111H, 4150 Clement St, San Francisco, CA 94121.*

*Dr Grimaldi's current address is Biostan Inc, 440 Chesapeake Dr, Seaport Centre, Redwood City, CA 94063.*

*The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked* "advertisement" *in accordance with 18 U.S.C. section 1734 solely to indicate this fact.*



*0006-4971/89/7308-0031$3.00/0*

lished methods.[5] Approximately 100 μg of high mol wt genomic DNA were partially digested with the *Sau*3A restriction enzyme. Fragments from 9 to 23 kilobases (kb) in size were isolated on a sucrose gradient and ligated into phage EMBL3A (Strategene, San Diego). Recombinant phage were packaged, plated, and screened as previously reported.[5]

*DNA sequencing.* Fragments for sequencing were cloned into M13 vectors and sequenced by the chain termination method using Sequenase (United States Biochemical, Cleveland).[7] All sequence data were derived from both strands.

## RESULTS

We studied a bone marrow sample from a patient with ALL and associated peripheral eosinophilia. Karyotypic analysis showed the characteristic t(5;14)(q31;q32) translocation. These features define a distinctive subtype of ALL.[3,4] The leukemic cells were analyzed for cell surface phenotype by immunofluorecence. They were positive for B1 (CD20), B4 (CD19), cALLA (CD10), HLA-DR, and terminal deoxynucleotidyl transferase (Tdt), but negative for surface immunoglobulin. This phenotypic profile describes an immature cell from the B-lymphocytic lineage.[8]

The leukemia DNA was analyzed by Southern blotting for rearrangements of the IgH gene. Using a human immunoglobulin Jh probe, a single rearranged band was detected by *Eco*RI, *Hind*III, *Sst*I, *Sau*3A, and *Eco*RI plus *Hind*III restriction digests, suggesting rearrangement of one allele (Fig 1). The immunoglobulin Jh region from the other allele was presumably either deleted or in the germline configuration.

We hypothesized that the t(5;14)(q31;q32) juxtaposed a growth-promoting gene on chromosome 5 with the immunoglobulin Jh region on chromosome 14. Therefore, a genomic library was made from the leukemic sample and screened with a Jh probe. Fifteen distinct positive clones were isolated and screened for the presence of the rearranged *Sau*3A fragment that was detected by DNA blotting. By this analysis, five clones appeared to represent the rearranged allele identified by DNA blots. One of these clones (clone no. 4) was chosen for further study and a detailed restriction map was generated. The *Eco*RI, *Hind*III/*Eco*RI, and *Sst*I fragments from clone no. 4 that hybridized to the human Jh probe were also identical in size to the rearranged fragments from the leukemia sample, confirming that clone no. 4 represented the rearranged leukemic allele.

Phage clone no. 4 contained 3.7 kb of unknown origin joined to the IgH gene in the region of Jh4 (Fig 2). The IgH gene from Jh4 to the Cmu region appeared to be in germline configuration. Previously, the gene encoding hematopoietic growth factor IL-3 had been mapped to chromosome 5q31 so it was suspected that clone no. 4 might contain part of this gene.[9-12] When the restriction map of human IL-3 and clone no. 4 were compared, they were identical for more than 3 kb (Fig 2).

We confirmed the juxtaposition of the IL-3 gene and the IgH gene by nucleic acid sequencing of the subcloned *Bst*EII/*Hpa*I fragment (Fig 2). The sequence of this fragment showed no disruption of the protein coding region or the messenger RNA of the IL-3 gene. The break in the IL-3 gene occurred in the promotor region, 452 base pairs (bp) upstream of the transcriptional start site (position 64, Fig

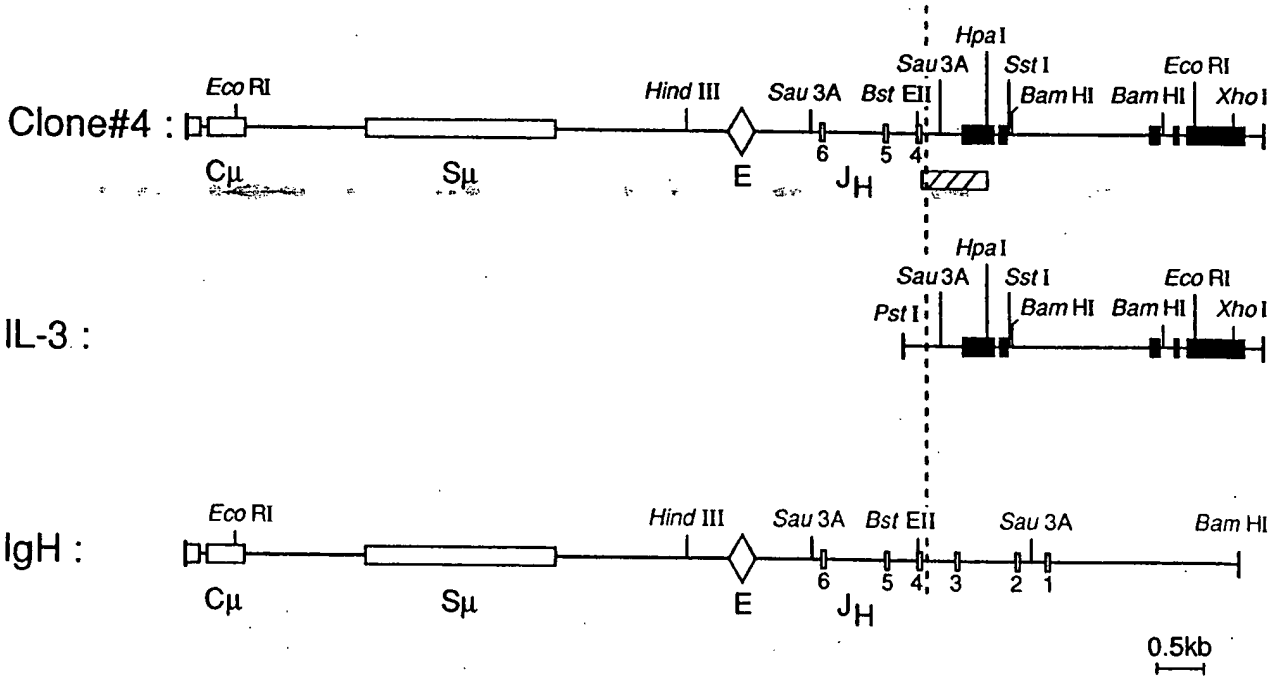



**Fig 2. Breakpoint region: t(5;14)(q31;q32). Comparative mapping of phage clone no. 4, the germline IgJh region, and the germline IL-3 gene.[28,29] The map of clone no. 4 is identical to that of IgH until it diverges in the region of Jh4 (at the dashed line), after which it is identical to the map of IL-3. The two genes are positioned in a head-to-head orientation. The Ig μ chain constant region (Cμ), switch region (Sμ), enhancer (E), and Jh segments are indicated (open symbols). The five exons (dark boxes) and four introns of the IL-3 gene are shown. The hatched box indicates the sequenced region.**

3A). The break in the IgH gene occurred 2 bp upstream of the Jh4 region. Between the two breaks, 25 bp of uncertain origin (putative N sequence) were inserted.[13,14] No sequences homologous to the immunoglobulin heptamer and nonamer could be identified in the IL-3 sequence (Fig 3B). Therefore, nucleic acid sequencing confirmed the juxtaposition of the IL-3 gene and the IgH gene. The sequence data clearly showed that the genes were positioned in opposite transcriptional orientations (head-to-head).

Available data also allowed us to determine the normal positions of the IL-3 gene and the GM-CSF gene in relation to the centromere of chromosome 5 (Fig 4). The IgH gene is known to be positioned with the variable regions toward the telomere on chromosome 14q.[2,15] It has also been shown that GM-CSF maps within 9 kb of IL-3 in the same transcriptional orientation.[16] Using this information and assuming a simple translocation event in our sample, we can conclude that the IL-3 gene is normally more centromeric, and the GM-CSF gene more telomeric on chromosome 5q (Fig 4). Furthermore, both are transcribed with their 5′ ends toward the centromere.

## DISCUSSION

In this report we have cloned a unique chromosomal translocation that appears to be a consistent feature of a rare, yet distinct, clinical form of acute leukemia. This translocation joined the promotor of the IL-3 gene to the IgH gene. Except for the altered promotor, the IL-3 gene appeared

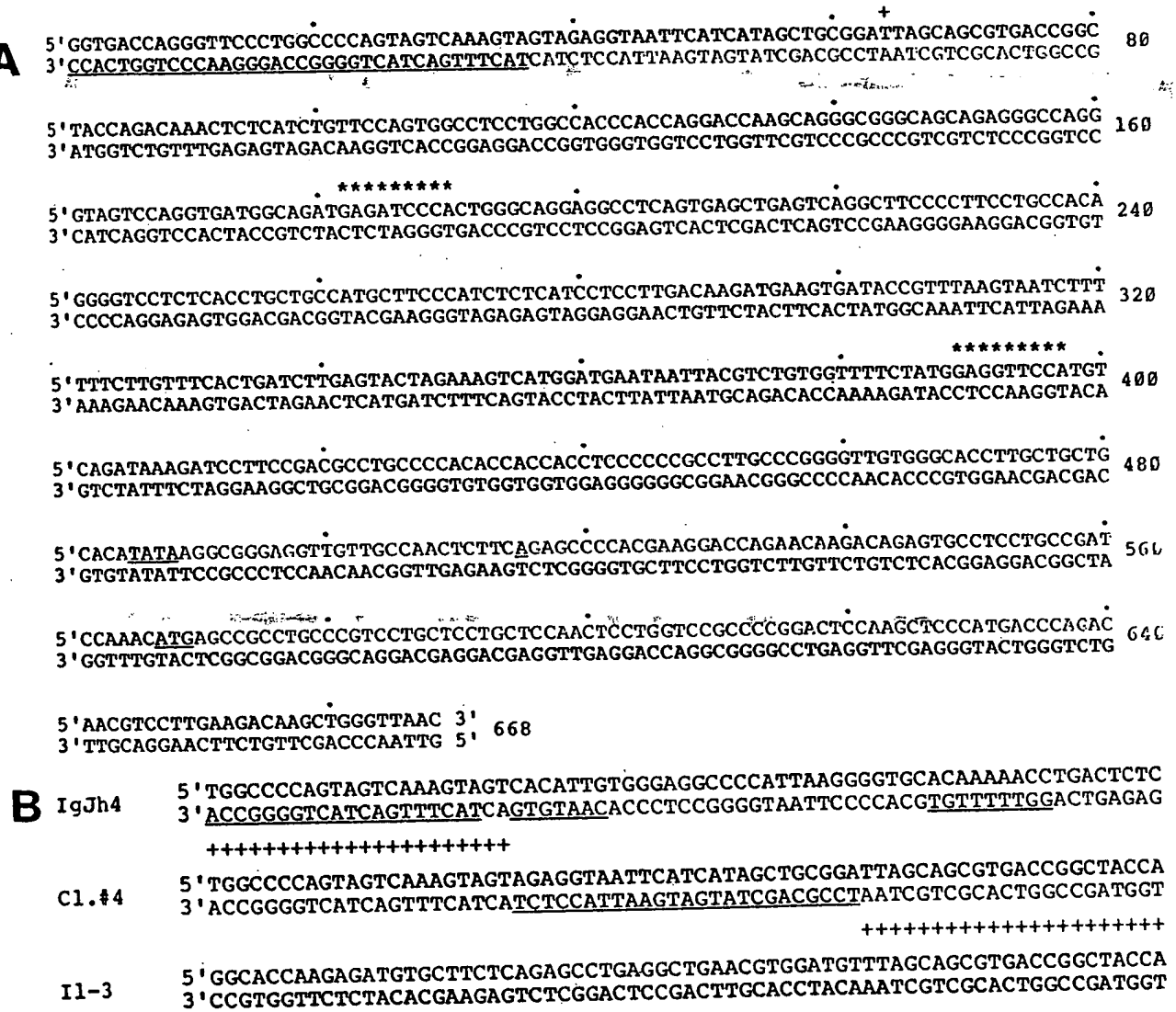



**Fig 3. Sequence of t(5;14)(q31;q32) breakpoint region. (A) Nucleotide sequence of the *Bst*EII/*Hpa*I fragment indicated on Fig 2. Nucleotides 1 to 36 represent the Jh4 coding region underlined on the coding strand.[9] Nucleotides 39 to 63 are a putative N region. The sequence from position 64 to 668 is that of the germline IL-3 gene.[28] The IL-3 TATA box (485), transcription start (515), and initiation methionine (567) are underlined. Two proposed regulatory sequences in the promotor are marked by asterisks (positions 182 and 389). (B) Comparative sequence of the t(5;14)(q31;q32) breakpoint region. The IgJh4 region is shown with its coding region, heptamer, and nonamer underlined. Clone no. 4 is shown with putative N region sequences underlined. The IL-3 sequence is also shown. A plus sign (+) denotes the identical nucleotide between sequences. No heptamer or nonamer is identified in the IL-3 sequence.**

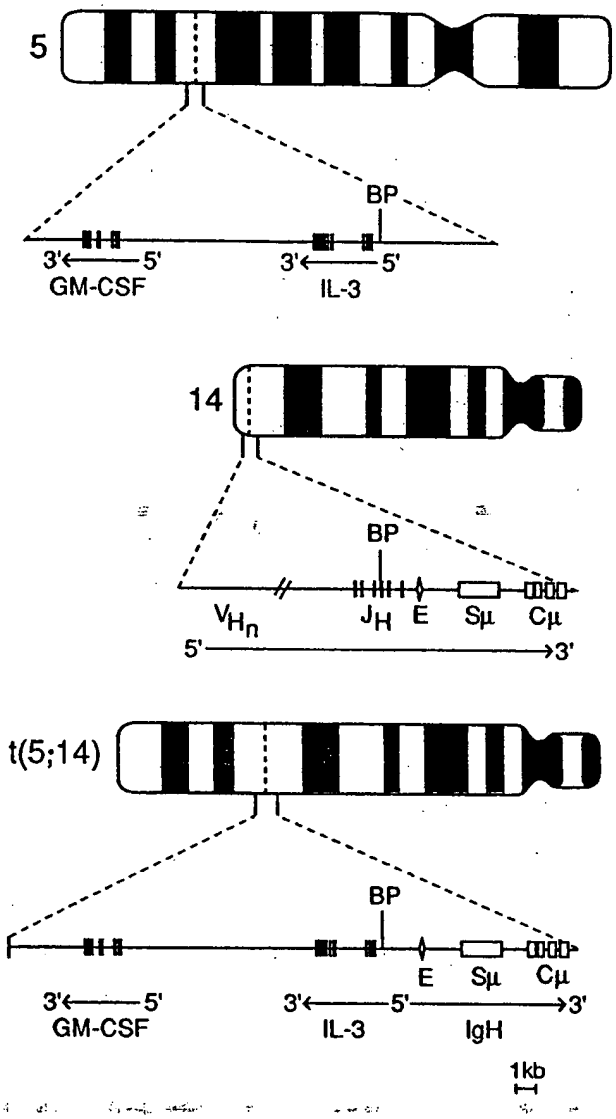



**Fig 4. Diagram of the translocation. The normal chromosome 5q31 is shown with the GM-CSF gene telomeric to the IL-3 gene in the transcriptional orientation shown. On normal chromosome 14q32 the Vh regions are telomeric. The t(5;14)(q31;q32) translocation results in the head-to-head orientation of these genes. Symbols are defined in Fig 2. BP, breakpoint position.**

intact as no deletions, insertions, or point mutations were detected by restriction mapping of the entire gene and sequencing of part of the gene. The IgH gene has been truncated at the Jh4 region, which places the immunoglobulin enhancer within 2.5 kb of the IL-3 gene.[17,18] This leads to the hypothesis that the enhancer is increasing transcription of a structurally normal IL-3 gene. The same mechanism is important for activation of the c-*myc* gene in some cases of Burkitt's lymphoma.[19] An alternate hypothesis is that the elimination of an upstream IL-3 promotor element is crucial to the activation of the IL-3 gene.

The proposed activation of the IL-3 gene suggests that an autocrine loop is important for the pathogenesis of this leukemia.[20] Over-expression of the IL-3 gene coupled with the presence of the IL-3 receptor in these cells could account for a strong stimulus for proliferation. In this regard, there are data indicating that immature B-lineage lymphocytes and B-lineage leukemias may express the IL-3 receptor.[21,22]

An additional feature of this type of leukemia is the dramatic eosinophilia, consisting of mature forms. It has been hypothesized that the eosinophils do not arise from the malignant clone, but are stimulated by the tumor.[23,24] Because of the known effect of IL-3 on eosinophil differentiation, secretion of high levels of IL-3 by leukemic cells might have a role in the eosinophilia in this type of leukemia.[12]

The data suggest that the recombination mechanism that is active in the IgH gene during normal differentiation has a role in this translocation.[13,14] This is supported by the breakpoint location at the 5′ end of Jh4 and the presence of putative N-region sequences. On the other hand, no recombination signal sequence (heptamer and nonamer) was found in this region on chromosome 5, suggesting that additional factors also played a role. Further studies will elucidate the mechanism of this and other translocations.

In the leukemia we studied, it is possible that the immunoglobulin enhancer also activates the GM-CSF gene, since this gene is probably positioned only 14 kb away (Fig 4). This is known to be within the range of enhancer activation.[25] The interleukin-5 (IL-5) gene maps to chromosome 5q31.[26] Deregulation of the IL-5 gene by this translocation would act synergistically with IL-3 in the stimulation of eosinophil proliferation and differentiation.[27] These and other questions will be answered by the study of more patient samples. We plan to determine whether the t(5;14)(q31;q32) translocation is capable of activating multiple lymphokines simultaneously and whether they cooperate in the generation of this leukemia.

## REFERENCES

1. Klein G, Klein E: Evolution of tumours and the impact of molecular oncology. Nature 315:190, 1985

2. Showe L, Croce C: The role of chromosomal translocations in B- and T-cell neoplasia. Annu Rev Immunol 5:253, 1987

3. Hogan T, Koss W, Murgo A, Amato R, Fontana J, VanScoy F: Acute lymphoblastic leukemia with chromosomal 5;14 translocation and hypereosinophilia: Case report and literature review. J Clin Oncol 5:382, 1987

4. Tono-oka T, Sato Y, Matsumoto T, Ueno N, Ohkawa M, Shikano T, Takeda T: Hypereosinophilic syndrome in acute lymphoblastic leukemia with a chromosome translocation t(5q;14q). Med Pediatr Oncol 12:33, 1984

5. Meeker T, Grimaldi JC, O'Rourke R, Loeb J, Juliusson G, Einhorn S: Lack of detectable somatic hypermutation in the V region of the IgH gene of a human chronic B-lymphocytic leukemia. J Immunol 141:3394, 1988

6. Ravetch J, Siebenlist U, Korsmeyer S, Waldmann T, Leder P: Structure of the human immunoglobulin μ locus: Characterization of embryonic and rearranged J and D genes. Cell 27:583, 1981

7. Norrander U, Kempe T, Messing J, Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene 26:101, 1983

8. Foon K, Todd R: Immunologic classification of leukemia and lymphoma. Blood 68:1, 1986

9. LeBeau M, Epstein N, O'Brien SJ, Nienhuis AW, Yang Y-C, Clark S, Rowley J: The interleukin-3 gene is located on human

chromosome 5 and is deleted in myeloid leukemias with a deletion of 5q. Proc Natl Acad Sci USA 84:5913, 1987

10. LeBeau M, Chandrasekharappi S, Lemons R, Schwartz J, Larson R, Arai N, Westbrook C: Molecular and cytogenetic analysis of chromosome 5 abnormalities in myeloid disorders, in cancer cells, in Proceedings of Molecular Diagnostics of Human Cancer. Cold Spring Harbor Laboratory, NY, 1989 (in press)

11. Ihle J, Weinstein Y: Immunological regulation of hematopoietic/lymphoid stem cell differentiation by interleukin-3. Adv Immunol 39:1, 1986

12. Clark S, Kamen R: The human hematopoietic colony-stimulating factors. Science 236:1229, 1987

13. Bakhshi A, Wright J, Graninger W, Seto M, Owens J, Cossman J, Jensen J, Goldman P, Korsmeyer S: Mechanism of the t(14,18) chromosomal translocation: Structural analysis of both derivative 14 and 18 reciprocal partners. Proc Natl Acad Sci USA 84:2396, 1987

14. Tsujimoto Y, Louie E, Bashir M, Croce C: The reciprocal partners of both the t(14,18) and the t(11,14) translocations involved in B-cell neoplasms are rearranged by the same mechanism. Oncogene 2:347, 1988

15. Erikson J, Finan J, Nowell P, Croce C: Translocation of immunoglobulin VH genes in Burkitt lymphoma. Proc Natl Acad Sci USA 80:810, 1982

16. Yang Y-C, Kovacic S, Kriz R, Wolf S, Clark S, Wellems T, Nienhuis A, Epstein N: The human genes for GM-CSF and IL-3 are closely linked in tandem on chromosome 5. Blood 71:958, 1988

17. Gillies S, Morrison S, Oi V, Tonegawa S: A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell 33:717, 1983

18. Banerji J, Olson L, Schaffner W: A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell 33:729, 1983

19. Hayday A, Gillies S, Saito H, Wood C, Wiman K, Hayward W, Tonegawa S: Activation of a translocated human c-*myc* gene by an enhancer in the immunoglobulin heavy-chain locus. Nature 307:334, 1984

20. Sporn M, Roberts A: Autocrine growth factors and cancer. Nature 313:745, 1985

21. Palacios R, Steinmetz M: IL-3–dependent mouse clones that express B-220 surface antigen, contain Ig genes in germ line configuration, and generate B lymphocytes in vivo. Cell 41:727, 1985

22. Uckun F, Gesner T, Song C, Myers D, Mufson A: Leukemic B-cell precursors express functional receptors for human interleukin-3. Blood 73:533, 1989

23. Spitzer G, Garson O: Lymphoblastic leukemia with marked eosinophilia: A report of two cases. Blood 42:377, 1973

24. Catovsky D, Bernasconi C, Verkonck P, Postma A, Howss J, Berg A, Rees J, Castelli G, Morra E, Galton D: The association of eosinophilia with lymphoblastic leukemia or lymphoma: A study of seven patients. Br J Haematol 45:523, 1980

25. Wang X-F, Calame K: The endogenous immunoglobulin heavy chain enhancer can activate tandem Vh promoters separated by a large distance. Cell 43:659, 1985

26. Sutherland G, Baker E, Callen D, Campbell H, Young I, Sanderson C, Garson O, Lopez A, Vadas M: Interleukin-5 is at 5q31 and is deleted in the 5q-syndrome. Blood 71:1150, 1988

27. Warren D, Moore M: Synergism among interleukin-1, interleukin-3, and interleukin-5 in the production of eosinophils from primitive hemopoietic stem cells. J Immunol 140:94, 1988

28. Yang Y-C, Clark S: Molecular cloning of a primate cDNA and the human gene for interleukin-3. Lymphokines 15:375, 1988

29. Yang Y-C, Ciarletta A, Temple P, Chung M, Kovacic S, Witek-Giannotti J, Leary A, Kriz R, Donahue R, Wong G, Clark S: Human IL-3 (multi-CSF): Identification by expression cloning of a novel hematopoietic growth factor related to murine IL-3. Cell 47:3, 1986

RAPID COMMUNICATION

# Activation of the Interleukin-3 Gene by Chromosome Translocation in Acute Lymphocytic Leukemia With Eosinophilia

By Timothy C. Meeker, Dan Hardy, Cheryl Willman, Thomas Hogan, and John Abrams

**The t(5;14)(q31;q32) translocation from B-lineage acute lymphocytic leukemia with eosinophilia has been cloned from two leukemia samples. In both cases, this translocation joined the IgH gene and the interleukin-3 (IL-3) gene. In one patient, excess IL-3 mRNA was produced by the leukemic cells. In the second patient, serum IL-3 levels were measured and shown to correlate with disease activity. There was no evidence of excess granulocyte/macrophage colony stimulating factor (GM-CSF) or IL-5 expression. Our data support the formulation that this subtype of leukemia may arise in part because of a chromosome translocation that activates the IL-3 gene, resulting in autocrine and paracrine growth effects.**


A NUMBER OF chromosome translocations have been associated with human leukemia and lymphoma. In many cases the study of these translocations has led to the discovery or characterization of proto-oncogenes, such as *bcl*-2, c-*abl*, and c-*myc*, that are located adjacent to the translocation.[1,2] It is now widely understood that cancer-associated translocations disrupt nearby proto-oncogenes.

A distinct subtype of acute leukemia is characterized by the triad of B-lineage immunophenotype, eosinophilia, and the t(5;14)(q31;q32) translocation.[3,4] Leukemic cells from such patients have been positive for terminal deoxynucleotidyl transferase (Tdt), common acute lymphoblastic leukemia antigen (CALLA), and CD19, but negative for surface or cytoplasmic immunoglobulin. In previous work, we cloned the t(5;14) breakpoint from one leukemic sample (Case 1) and determined that the IgH and interleukin-3 (IL-3) genes were joined by this abnormality.[5] In this report, we extend those findings by showing that the t(5;14)(q31;q32) translocation from a second leukemia sample (Case 2) has a similar structure, and we report our study of growth factor expression in these patients.

## MATERIALS AND METHODS

*Samples and Southern blots.* Case 1 has been described.[5,6] Clinical features of Case 2 have been described in detail.[3] DNA isolation and Southern blotting was done using previously described methods.[5] Filters were hybridized with an immunoglobulin Jh probe, a 280 bp *Bam*HI/*Eco*RI genomic IL-3 fragment, and an IL-3 cDNA probe.[7,8]

*Northern blots.* RNA isolation and Northern blotting have been described.[9] Briefly, Northern blots were done by separating 9μg total RNA on 1% agarose-formaldehyde gels. Equal RNA loading in each lane was confirmed by ethidium bromide staining. Blots were hybridized with an IL-3 cDNA probe extending to the *Xho* I site in exon 5, a 720 bp *Sst* I/*Kpn* I probe derived from intron 2 of the IL-3 gene, a 600 bp *Nhe* I/*Hpa* I IL-5 cDNA probe, and a 500 bp *Pst* I/*Nco* I granulocyte-macrophage colony stimulating factor (GM-CSF) cDNA probe.[10-12]

*Polymerase chain reaction.* Primers were designed with *Bam*HI sites for cloning. One primer hybridized to the Jh sequences from the IgH gene (Primer 144:5'-TAGGATCCGACGGTGACCAGGGT), and the other hybridized to the region of the TATA box in the IL-3 gene (Primer 161: 5'-AACAGGATCCCGCCTTATATGTGCAG). Polymerase chain reaction (PCR) (95°C for 1 minute, 61°C for 30 seconds, and 72°C for 3 minutes) was done using 500 ng genomic DNA and 50 pmol of each primer in 100 μL containing 67 mmol/L Tris-HCl pH 8.8, 6.7 mmol/L $MgCl_2$, 10% dimethyl sulfoxide (DMSO), 170 μg/mL bovine serum albumin (BSA) (fraction V), 16.6 mmol/L ammonium sulfate, 1.5 mmol/L each dNTP and Taq polymerase (Perkin-Elmer, Norwalk, CT).[13]

*Sequencing.* Sequencing was done by chain termination in M13 vectors.[14] As part of this study, we sequenced a subclone of a normal IL-3 promotor, covering 598 base pairs from a *Sma* I site at position −1240 (with respect to the proposed site of transcription initiation) to an *Nhe* I site at position −642. The plasmid containing this region was a gift from Naoko Arai of the DNAX Research Institute.

*Expression in Cos7 cells.* A genomic IL-3 fragment from Case 1 was cloned into the pXM expression vector.[10] Briefly, the *Hin*dIII/*Sal* I fragment containing the IL-3 gene was subcloned from the previously described phage clone 4 into pUC18.[5] The 2.6 kb fragment extending from the *Sma* I site 61 bp upstream of the IL-3 transcription start to the *Sma* I site in the polylinker was cloned into the blunted *Xho* I site of pXM. The negative control construct was the pXM vector without insert. Plasmids were introduced into Cos7 cells by electroporation, and supernatant was collected after 48 hours in culture.

*TF1 bioassay.* TF-1 cells were passaged in RPMI 1640 supplemented with 10% heat-inactivated fetal bovine serum, 2 mmol L-glutamine, and 1 ng/mL human GM-CSF.[15] Samples and antibodies were diluted in this same medium lacking GM-CSF but containing penicillin and streptomycin. A 25 μL volume of serial dilutions of patient serum was added to wells in a flat bottom 96-well microtiter plate. Rat anti-cytokine monoclonal antibody in a volume of 25 μL was added to appropriate wells and preincubated for 1 hour at 37°C. Fifty microliters of twice washed TF-1 cells were added to each well, giving a final cell concentration of $1 \times 10^4$ cells per well (final volume, 100 μL). The plate was incubated for 48 hours. The remaining cell viability was determined metabolically by the colori-

*From the Division of Hematology/Oncology 111H, Department of Medicine, University of California and the Veterans Administration Medical Center, San Francisco, CA; the Center for Molecular and Cellular Diagnostics, Department of Pathology and Cell Biology, University of New Mexico, Albuquerque, NM; the Division of Hematology/Oncology, Department of Medicine, West Virginia University, Morgantown, WV; and DNAX Research Institute, Palo Alto, CA.*

*Submitted March 27, 1990; accepted April 19, 1990.*

*Supported in part by the University of California Cancer Research Coordinating Committee and University of New Mexico Cancer Center funding from the state of New Mexico. The DNAX Research Institute is supported by Schering-Plough.*

*Address reprint requests to Timothy C. Meeker, MD, Division of Hematology/Oncology 111H, Department of Medicine, University of California and the Veterans Administration Medical Center, 4150 Clement St, San Francisco, CA 94121.*



*0006-4971/90/7602-0022$3.00/0*

```
IgJh5   5'GACGGTGACCAGGGTTCCTTGGCCCCAGGAGTCGAACCAGTTGTCACATTGTGACAACAATGCCAGACCCCGACAAAGAACGGGCGCCCCCGCACGTCCCT
        3'CTGCCACTGGTCCCAAGGAACCGGGGTCCTCAGCTTGGTCAACAGTGTAACACTGTTGTTACGGTCTGGGGCTGTTTCTTGCCCGCGGGGCGTGCAGGGA
          ++++++++++++++++++ ++++++++++ ++++++++++++++

Clone   5'GACGGTGACCAGGGTTCCCTGGCCCCAGGGGTCGAACCAGTTGTACCGCCATGTTATTGGGGGGGATCAAGACCTCTCAATACAACTGTCTCCTGCCAAT
        3'CTGCCACTGGTCCCAAGGGACCGGGGTCCCCAGCTTGGTCAACATGGCGGTACAATAACCCCCCCTAGTTCTGGAGAGTTATGTTGACAGAGGACGGTTA
                                                                    (-934)*++++++++++++++++++++++++++++++++++++++++

IL3     5'CCCCTCTCTGCAAACCTTGCCTACTGGGCCTGCACCTGGCAAATCCATGCTCAGCACAGACGGGGATCAAGACCTCTCAATACAACTGTCTCCTGCCAAT
        3'GGGGAGAGACGTTTGGAACGGATGACCCGGACGTGGACCGTTTAGGTACGAGTCGTGTCTGCCCCTAGTTCTGGAGAGTTATGTTGACAGAGGACGGTTA
```

**Fig 1. Breakpoint sequences for Case 2. The germline IgJh5 region sequence (protein coding region and recombination signal sequences are underlined) is on top, the translocation sequence from Case 2 (PCR primer sequences and putative N region are underlined) is in the middle, and the germline IL-3 sequence, which we derived from a normal IL-3 clone, is on the bottom.[7] + indicates that each sequence has the same nucleotide. The sequence documents the head-to-head joining of the IL-3 and IgH genes. The breakpoint in the IL-3 gene occurred at position −934 (*).**

metric method of Mosmann using a VMax microtiter plate reader (Molecular Devices, Menlo Park, CA) set at 570 and 650 nm.[16]

*Cytokine immunoassays.* These assays used rat monoclonal anti-cytokine antibodies (10 μg/mL) to coat the wells of a PVC microtiter plate. The capture antibodies used were BVD3-6G8, JES1-39D10, and BVD2-23B6, for the IL-3, IL-5, and GM-CSF assays, respectively. Patient sera were then added (undiluted and diluted 1:2 for IL-3, undiluted for IL-5, and undiluted and diluted 1:5 for GM-CSF). The detecting immunoreagents used were either mouse antiserum to IL-3 or nitroiodophenyl (NIP)-derivatized rat monoclonal antibodies JES1-5A2 and BVD2-21C11, specific for IL-5 and GM-CSF, respectively. Bound antibody was subsequently detected with immunoperoxidase conjugates: horseradish peroxidase (HRP)-labeled goat anti-mouse Ig for IL-3, or HRP-labeled rat (J4 MoAb) anti-NIP for IL-5 and GM-CSF. The chromogenic substrate was 3-3'azino-bis-benzthiazoline sulfonate (ABTS; Sigma, St Louis, MO). Unknown values were interpolated from standard curves prepared from dilutions of the recombinant factors using Softmax software available with the VMAX microplate reader (Molecular Devices).

## RESULTS

Leukemic DNA from Case 2 was studied by Southern blotting. When digested with the *Hin*dIII restriction enzyme and hybridized with a human immunoglobulin heavy chain joining region (Jh) probe, a rearranged fragment at approximately 14 kb was detected (data not shown). When reprobed with either of two different IL-3 probes, a rearranged 14 kb fragment, comigrating with the rearranged Jh fragment, was identified. When leukemic DNA was digested with *Hin*dIII plus *Eco*RI, a rearranged Jh fragment was detected at 6 kb. The IL-3 probes also identified a comigrating fragment of this size. These experiments indicated that the leukemic sample studied was clonal and that a single fragment contained both Jh and IL-3 sequences, suggesting a translocation had occurred.

To characterize better the joining of the IL-3 gene and the immunoglobulin heavy chain (IgH) gene, the polymerase chain reaction (PCR) was used to clone the translocation.[13] A Jh primer and an IL-3 primer were designed to produce an amplified product in the event of a head-to-head translocation. While control DNA gave no PCR product, Case 2 DNA yielded a PCR-derived fragment of approximately 980 bp, which was cloned and sequenced.

The DNA sequence of the translocation clone from Case 2 confirmed the joining of the Jh region with the promotor of the IL-3 gene in a head-to-head configuration (Fig 1). Sequence analysis indicated that the breakpoint on chromosome 14 was just upstream of the Jh5 coding region. The breakpoint on chromosome 5 occurred 934 bp upstream of the putative site of transcription initiation of the IL-3 gene. We also determined that a putative N sequence of 17 bp was inserted between the chromosome 5 and chromosome 14 sequences during the translocation event.[17,18] Figure 2 shows

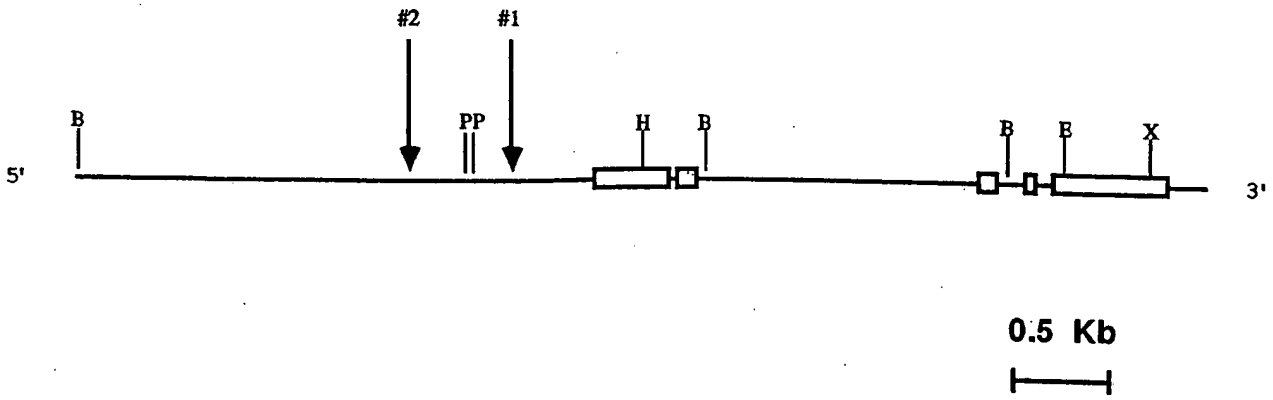



**Fig 2. Relationship of chromosome 5 breakpoints to the IL-3 gene. This figure shows the two cloned breakpoints (arrows) in relation to the normal IL-3 gene.[9,10] One breakpoint occurred at position −462 and the other at −934 (arrows). In both circumstances, the translocations resulted in a head-to-head joining of the IgH gene and the IL-3 gene, leaving the mRNA and protein coding regions of the IL-3 gene intact. Boxes denote the five IL-3 exons; restriction enzymes are (B) *Bam*HI, (P) *Pst* I, (H) *Hpa* I, (E) *Eco*RI, and (X) *Xho* I.**

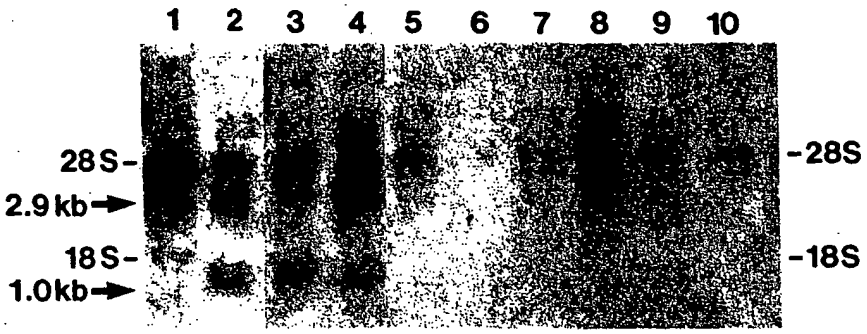



**Fig 3. Documentation of IL-3 mRNA over-expression. A Northern blot was prepared and hybridized with a probe for IL-3. Lane 1 contained RNA from unstimulated peripheral blood lymphocytes (PBL) as a negative control. Lane 2 contained RNA from PBL stimulated for 4 hours with concanavalian A (ConA), and lane 3 contained RNA from PBL stimulated with ConA for 48 hours. As in the positive control lanes (2 and 3), a 1 kb band was identified in the leukemic sample from Case 1 (lane 4, lower arrow), suggesting aberrant expression of the IL-3 gene. In addition, the leukemic sample showed over-expression of an unspliced 2.9 kb IL-3 transcript (lane 4, upper arrow). We documented that this represented an unspliced precurser of the mature 1 kb transcript by showing that this band hybridized to a probe from intron 2 of the IL-3 gene. A similar 2.9 kb band was detected in lane 2, suggesting that an IL-3 mRNA of this size is sometimes detectable in normal mitogen-stimulated cells. Lane 5 through 10 represent RNA from six samples of B-lineage acute lymphocytic leukemia without the t(5;14) translocation, indicating that only the sample with the translocation exhibited IL-3 over-expression. Case 2 could not be analyzed by Northern blot because too few cells were available for study.**

the locations of the two cloned breakpoints in relation to the IL-3 gene. The two chromosome 5 breakpoints were separated by less than 500 bp.

The genomic structure in Cases 1 and 2 suggested that a normal IL-3 gene product was over-expressed as a result of the altered promotor structure. This would predict that the IL-3 gene on the translocated chromosome was capable of making IL-3 protein. This prediction was tested by expressing a genomic fragment from the translocated allele of Case 1 containing all five IL-3 exons under the control of the SV40 promotor/enhancer in the Cos7 cell line. Cell supernatants were studied in a proliferation assay using the factor dependent erythroleukemic cell line, TF-1. The supernatants derived from transfections using the vector plus insert supported TF-1 proliferation, while supernatants from transfections using the vector alone were negative in this assay (data not shown). Furthermore, the biologic activity could be blocked by an antibody to human IL-3 (BVD3-6G8). This result showed that the translocated allele retained the ability to make IL-3 mRNA and protein.

The level of expression of IL-3 mRNA in leukemic cells from Case 1 was assessed. Northern blotting showed that the mature IL-3 mRNA (approximately 1 kb) and a 2.9 kb unspliced IL-3 mRNA were excessively produced by the leukemia (Fig 3). The 2.9 kb form of the mRNA is also present at low levels in normal peripheral blood T lymphocytes after mitogen activation (Fig 3). Several B-lineage acute leukemia samples without the t(5;14) translocation had undetectable levels of IL-3 mRNA in these experiments. In addition, although genes for GM-CSF and IL-5 map close to the IL-3 gene and might have been deregulated by the translocation, no IL-5 or GM-CSF mRNA could be detected in the leukemic sample (data not shown).[19,20]

Three serum samples from Case 2 were assayed by immunoassay for levels of IL-3, GM-CSF, and IL-5 (Table 1). Serum IL-3 could be detected and correlated with the clinical course. When the patient's leukemic cell burden was highest, the IL-3 level was highest. No serum GM-CSF or IL-5 could be detected.

Since the IL-3 immunoassay measured only immunoreactive factor, we confimed that biologically active IL-3 was present by using the TF-1 bioassay. This bioassay can be rendered monospecific using appropriate neutralizing monoclonal antibodies specific for IL-3, IL-5, or GM-CSF. We observed that sera from 1-16-84 and 3-14-84 contained TF-1 stimulating activity that could be blocked with anti–IL-3 MoAb (BVD3-6G8), but not with MoAbs to IL-5 (JES1-39D10) or GM-CSF (BVD2-23B6) (Fig 4; GM-CSF data not shown). The amount of neutralizable bioactivity in these two samples correlated very well with the difference in IL-3 levels obtained by immunoassay for these samples. Furthermore, the failure to block TF-1 proliferating activity with either anti–IL-5 or anti–GM-CSF was consistent with the inability to measure these factors by immunoassay and

**Table 1. Peripheral Blood Counts and Growth Factor Levels at Different Times in Case 2**

| | Sample Date | | |
|---|---|---|---|
| | 11/15/83 | 1/16/84 | 3/14/84 |
| Peripheral blood counts (cells/μL) | | | |
| WBC | 81,800 | 116,500 | 12,300 |
| Lymphoblasts | 0 | 33,785 | 0 |
| Eosinophils | 46,626 | 73,080 | 615 |
| Serum growth factor levels (pg/mL) | | | |
| IL-3 | <444 | 7,995 | 1,051 |
| GM-CSF | <15 | <15 | <15 |
| IL-5 | <50 | <50 | <50 |

Peripheral blood counts from Case 2 at three different time points with the corresponding growth factor levels quantified by immunoassay. The patient received chemotherapy between 1/16/84 and 3/14/84 to lower his leukemic burden.[3] No serum samples were available for a similar analysis of Case 1.

Abbreviation: WBC, white blood cells.

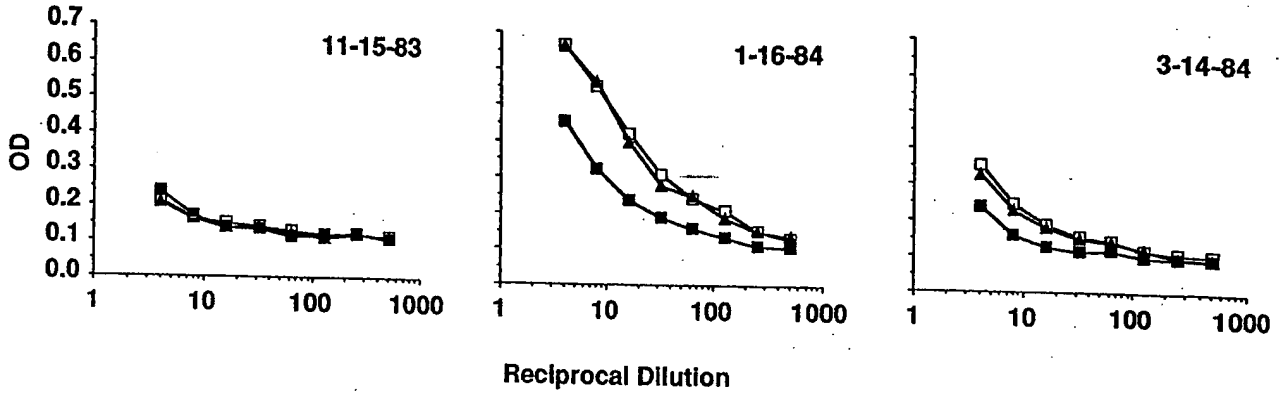



**Fig 4. Bioassay of serum IL-3. Leukemic patient sera were tested for bioactive IL-3 and IL-5 in the TF-1 proliferation assay. The reciprocal of the dilution is indicated on the horizontal axis and the optical density indicating the amount of proliferation is indicated on the vertical axis. Serum from all three time points was assayed simultaneously. The assay was rendered monospecific by using a 1 μg/mL final concentration of monoclonal rat anti–IL-3, BVD3-6G8 (■), or anti–IL-5, JES1-39D10 (▲); □ indicates no MoAb. On 1/16/84 and 3/14/84, inhibition of proliferation was evident in the presence of anti–IL-3 antibody, documenting serum levels of IL-3 on those days. Serum IL-5 was not detected in this assay, as anti–IL-5 did not alter TF-1 proliferation.**

indicated that these other myeloid growth factors were not detectably circulating in the serum of this patient.

## DISCUSSION

In this report, we have extended our analysis of acute lymphocytic leukemia and eosinophilia associated with the t(5;14) translocation. In both cases we have studied, we have documented the joining of the IL-3 gene from chromosome 5 to the IgH gene from chromosome 14. The breakpoints on chromosome 5 are within 500 bp of each other, suggesting that additional breakpoints will be clustered in a small region of the IL-3 promotor. The PCR assay we have developed will be useful in the screening of additional clinical samples for this abnormality.

The finding of a disrupted IL-3 promotor associated with an otherwise normal IL-3 gene implied that this translocation might lead to the over-expression of a normal IL-3 gene product. In this work, we have documented that this is true. In addition, neither GM-CSF nor IL-5 are over-expressed by the leukemic cells. Furthermore, in one patient, serum IL-3 could be measured and correlated with disease activity. To our knowledge, this is the first measurement of human IL-3 in serum and its association with a disease process. The measurement of serum IL-3 in this and other clinical settings may now be indicated.

The finding of the IL-3 gene adjacent to a cancer-associated translocation breakpoint suggests that its activation is important for oncogenesis. It is our thesis that an autocrine loop for IL-3 is important for the evolution of this leukemia.[21] The excessive IL-3 production that we have documented would be one feature of such an autocrine loop. The final proof of our thesis must await additional data. In particular, from the study of additional clinical samples, it will be necessary to document that the IL-3 receptor is present on the leukemic cells and that anti–IL-3 antibody decreases proliferation of the leukemia in vitro.

An important aspect of this work is the suggestion of a therapeutic approach for this disease. If an autocrine loop for IL-3 can be documented in this disease, attempts to lower circulating IL-3 levels or block the interaction of IL-3 with its receptor may prove useful. Because it is also possible that the eosinophilia in these patients is mediated by the paracrine effects of leukemia-derived IL-3, similar interventions may improve this aspect of the disease. Antibodies or engineered ligands to accomplish these goals may soon be available.

## ACKNOWLEDGMENT

We thank Naoko Arai, Ken-ichi Arai, R. O'Rourke, J. Grimaldi, and T. O'Connell for technical assistance and/or helpful discussions.

## REFERENCES

1. Klein G, Klein E: Evolution of tumours and the impact of molecular oncology. Nature 315:190, 1985

2. Showe L, Croce C: The role of chromosomal translocations in B- and T-cell neoplasia. Ann Rev Immunol 5:253, 1987

3. Hogan T, Koss W, Murgo A, Amato R, Fontana J, VanScoy F: Acute lymphoblastic leukemia with chromosomal 5;14 translocation and hypereosinophilia: Case report and literature review. J Clin Oncol 5:382, 1987

4. Tono-oka T, Sato Y, Matsumoto T, Ueno N, Ohkawa M, Shikano T, Takeda T: Hypereosinophilic syndrome in acute lymphoblastic leukemia with a chromosome translocation t(5q;14q). Med Ped Oncol 12:33, 1984

5. Grimaldi J, Meeker T: The t(5;14) chromosomal translocation in a case of acute lymphocytic leukemia joins the interleukin-3 gene to the immunoglobulin heavy chain gene. Blood 73:2081, 1989

6. McConnell T, Foucar K, Hardy W, Saiki J: Three-way reciprocal chromosomal translocation in an acute lymphoblastic leukemia patient with hypereosinophilia syndrome. J Clin Oncol 5:2042, 1987

7. Ravetch J, Siebenlist U, Korsmeyer S, Waldmann T, Leder P: Structure of the human immunoglobulin m locus: Characterization of embryonic and rearranged J and D genes. Cell 27:583, 1981

8. Otsuka T, Miyajima A, Brown N, Otsu K, Abrams J, Saeland S, Caux C, Malefijt R, Vries J, Meyerson P, Yokota K, Gemmel L,

Rennick D, Lee F, Arai N, Arai K, Yokota T: Isolation and characterization of an expressible cDNA encoding human IL-3. J Immunol 140:2288, 1988

9. Sambrook J, Fritsch E, Maniatis T: Molecular Cloning. Cold Spring Harbor, NY, Cold Spring Harbor Press, 1989

10. Yang Y-C, Ciarletta A, Temple P, Chung M, Kovacic S, Witek-Giannotti J, Leary A, Kriz R, Donahue R, Wong G, Clark S: Human IL-3 (multi-CSF): Identification by expression cloning of a novel hematopoietic growth factor related to murine IL-3. Cell 47:3, 1986

11. Yokota T, Coffman R, Hagiwara H, Rennick D, Takebe Y, Yokota K, Gemmell L, Shrader B, Yang G, Meyerson P, Luh J, Hoy P, Pene J, Briere F, Spits H, Banchereau J, Vries J, Lee F, Arai N, Arai K: Isolation and characterization of lymphokine cDNA clones encoding mouse and human IgA-enhancing factor and eosinophil colony-stimulating factor activities: Relationship to interleukin 5. Proc Natl Acad Sci USA 84:7388, 1987

12. Wong G, Witek J, Temple P, Wilkens K, Leary A, Luxenberg D, Jones S, Brown E, Kay R, Orr E, Shoemaker C, Golde D, Kaufman R, Hewick R, Wang E, Clark S: Human GM-CSF: Molecular cloning of the complementary DNA and purification of the natural and recombinant proteins. Science 228:810, 1985

13. Saiki R, Scharf S, Faloona F, Mullis K, Horn G, Erlich H, Arnheim N: Enzymatic amplification of B-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230:1350, 1985

14. Norrander U, Kempe T, Messing J: Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene 26:101, 1983

15. Kitamura T, Tange T, Terasawa T, Chiba S, Kuwaki T, Miyagawa K, Piao Y, Miyazono K, Urabe A, Takaku F: Establishment and characterization of a unique human cell line that proliferates dependently on GM-CSF, IL-3, or erythropoietin. J Cell Physiol 140:323, 1989

16. Mosmann T: Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. J Immunol Methods 65:55, 1983

17. Bakhshi A, Wright J, Graninger W, Seto M, Owens J, Cossman J, Jensen J, Goldman P, Korsmeyer S: Mechanism of the t(14;18) chromosomal translocation: Structural analysis of both derivative 14 and 18 reciprocal partners. Proc Natl Acad Sci USA 84:2396, 1987

18. Tsujimoto Y, Louie E, Bashir M, Croce C: The reciprocal partners of both the t(14;18) and the t(11;14) translocations involved in B-cell neoplasms are rearranged by the same mechanism. Oncogene 2:347, 1988

19. Yang Y-C, Kovacic S, Kriz R, Wolf S, Clark S, Wellems T, Nienhuis A, Epstein N: The human genes for GM-CSF and IL-3 are closely linked in tandem on chromosome 5. Blood 71:958, 1988

20. Sutherland G, Baker E, Callen D, Campbell H, Young I, Sanderson C, Garson O, Lopez A, Vadas M: Interleukin-5 is at 5q31 and is deleted in the 5q− syndrome. Blood 71:1150, 1988

21. Sporn M, Roberts A: Autocrine growth factors and cancer. Nature 313:745, 1985

# Clinical and Pathologic Significance of the c-*erb*B-2 (*HER*-2/*neu*) Oncogene

Timothy P. Singleton and John G. Strickler

The c-*erb*B-2 oncogene was first shown to have clinical significance in 1987 by Slamon et al,[79] who reported that c-*erb*B-2 DNA amplification in breast carcinomas correlated with decreased survival in patients with metastasis to axillary lymph nodes. Subsequent studies, however, of c-*erb*B-2 activation in breast carcinoma reached conflicting conclusions about its clinical significance. This oncogene also has been reported to have clinical and pathologic implications in other neoplasms. Our review summarizes these various studies and examines the clinical relevance of c-*erb*B-2 activation, which has not been emphasized in recent reviews.[37,38,55] The molecular biology of the c-*erb*B-2 oncogene has been extensively reviewed[37,38,55] and will be discussed only briefly here.

## BACKGROUND

The c-*erb*B-2 oncogene was discovered in the 1980s by three lines of investigation. The *neu* oncogene was detected as a mutated transforming gene in neuroblastomas induced by ethylnitrosurea treatment of fetal rats.[8,73,74,78] The c-*erb*B-2 was a human gene discovered by its homology to the retroviral gene v-*erb*B.[33,49,76] *HER*-2 was isolated by screening a human genomic DNA library for homology with v-*erb*B.[24] When the DNA sequences were determined subsequently, c-*erb*B-2, *HER*-2, and *neu* were found to represent the same gene. Recently, the c-*erb*B-2 oncogene also has been referred to as *NGL*.

The c-*erb*B-2 DNA is located on human chromosome 17q21[24,33,66] and codes for c-*erb*B-2 mRNA (4.6 kb), which translates c-*erb*B-2 protein (p185). This

protein is a normal component of cytoplasmic membranes. The c-*erb*B-2 oncogene is homologous with, but not identical to, c-*erb*B-1, which is located on chromosome 7 and codes for the epidermal growth factor receptor.[8,103] The c-*erb*B-2 protein is a receptor on cell membranes and has intracellular tyrosine kinase activity and an extracellular binding domain.[2,105] Electron microscopy with a polyclonal antibody detects c-*erb*B-2 immunoreactivity on cytoplasmic membranes of neoplasms, especially on microvilli and the non-villous outer cell membrane.[61] In normal cells, immunohistochemical reactivity for c-*erb*B-2 is frequently present at the basolateral membrane or the cytoplasmic membrane's brush border.[22,62]

There is experimental evidence that c-*erb*B-2 protein may be involved in the pathogenesis of breast neoplasia. Overproduction of otherwise normal c-*erb*B-2 protein can transform a cell line into a malignant phenotype.[25] Also, when the *neu* oncogene containing an activating point mutation is placed in transgenic mice with a strong promoter for increased expression, the mice develop multiple independent mammary adenocarcinomas.[16,63] In other experiments, monoclonal antibodies against the *neu* protein inhibit the growth (in nude mice) of a *neu*-transformed cell line,[26–28] and immunization of mice with *neu* protein protects them from subsequent tumor challenge with the *neu*-transformed cell line.[14] Some authors have speculated that the use of antagonists for the unknown ligand could be useful in future chemotherapy.[85] Further review of this experimental evidence is beyond the scope of this article.

The c-*erb*B-2 activation most likely occurs at an early stage of neoplastic development. This hypothesis is supported by the presence of c-*erb*B-2 activation in both in situ and invasive breast carcinomas. In addition, studies of metastatic breast carcinomas usually demonstrate uniform c-*erb*B-2 activation at multiple sites in the same patient,[11,12,39,41,52] although c-*erb*B-2 activation has rarely been detected in metastatic lesions but not in the primary tumor.[67,69,107] Even more rarely, c-*erb*B-2 DNA amplification has been detected in a primary breast carcinoma but not in its lymph node metastasis.[5] In patients who have bilateral breast neoplasms, both lesions have similar patterns of c-*erb*B-2 activation, but only a few such cases have been studied.[11]

## MECHANISMS OF c-*erb*B-2 ACTIVATION

The most common mechanism of c-*erb*B-2 activation is genomic DNA amplification, which almost always results in overproduction of c-*erb*B-2 mRNA and protein.[17,34,65,81] The c-*erb*B-2 amplification may stabilize the overproduction of mRNA or protein through unknown mechanisms. Human breast carcinomas with c-*erb*B-2 amplification contain 2 to 40 times more c-*erb*B-2 DNA[4,5] and 4 to 128 times more c-*erb*B-2 mRNA[34,90] than found in normal tissue. Most human breast carcinomas with c-*erb*B-2 amplification have 2 to 15 times more c-*erb*B-2 DNA. Tumors with greater amplification tend to have greater overproduction.[17,52,65] The non-mammary neoplasms that have been studied tend to have

similar levels of c-*erb*B-2 amplification or overproduction relative to the corresponding normal tissue.

The second most common mechanism of c-*erb*B-2 activation is overproduction of c-*erb*B-2 mRNA and protein without amplification of c-*erb*B-2 DNA.[81] The quantities of mRNA and protein usually are less than those in amplified cases and may approach the small quantities present in normal breast or other tissues.[17,50,52] The c-*erb*B-2 protein overproduction without mRNA overproduction or DNA amplification has been described in a few human breast carcinoma cell lines.[47]

Other rare mechanisms of c-*erb*B-2 activation have been reported. Translocations involving the c-*erb*B-2 gene have been described in a few mammary and gastric carcinomas, although some reported cases may represent restriction fragment length polymorphisms or incomplete restriction enzyme digestions that mimic translocations.[31,65,78,84,90,108] A single point mutation in the transmembrane portion of *neu* has been described in rat neuroblastomas induced by ethylnitrosurea.[9,55] The mutated *neu* protein has increased tyrosine kinase activity and aggregates at the cell membrane.[10,83,98] Although there has been speculation that some of the amplified c-*erb*B-2 genes may contain point mutations,[46] none has been detected in primary human neoplasms.[41,53,61]

## TECHNIQUES FOR DETECTING c-*erb*B-2 ACTIVATION

### Detection of c-*erb*B-2 DNA Amplification

Amplification of c-*erb*B-2 DNA is usually detected by DNA dot blot or Southern blot hybridization. In the dot blot method, the extracted DNA is placed directly on a nylon membrane and hybridized with a c-*erb*B-2 DNA probe. In the Southern blot method, the extracted DNA is treated with a restriction enzyme, and the fragments are separated by electrophoresis, transferred to a nylon membrane, and hybridized with a c-*erb*B-2 DNA probe. In both techniques, c-*erb*B-2 amplification is quantified by comparing the intensity (measured by densitometry) of the hybridization bands from the sample with those from control tissue.

Several technical problems may complicate the measurement of c-*erb*B-2 DNA amplification. First, the extracted tumor DNA may be excessively degraded or diluted by DNA from stromal cells.[61] Second, the c-*erb*B-2 DNA probe must be carefully chosen and labeled. For example, oligonucleotide c-*erb*B-2 probes may not be sensitive enough for measuring a low level of c-*erb*B-2 amplification, because diploid copy numbers can be difficult to detect (unpublished data). Third, the total amounts of DNA in the sample and control tissue must be compensated for, often with a probe to an unamplified gene. Many studies have used control probes to genes on chromosome 17, the location of c-*erb*B-2, to correct for possible alterations in chromosome number. Identical results, however, are obtained by using control probes to genes on other chromosomes,[5,65,80] with rare exception.[17] Studies using control probes to the beta-

globin gene must be interpreted with caution, because one allele of this gene is deleted occasionally in breast carcinomas.[3]

Amplification of c-*erb*B-2 DNA was assessed by using the polymerase chain reaction (PCR) in one recent study.[32] Oligoprimers for the c-*erb*B-2 gene and a control gene are added to the sample's DNA, and PCR is performed. If the sample contains more copies of c-*erb*B-2 DNA than of the control gene, the c-*erb*B-2 DNA is replicated preferentially.

## Detection of c-*erb*B-2 mRNA Overproduction

Overproduction of c-*erb*B-2 mRNA usually is measured by RNA dot blot or Northern blot hybridization. Both techniques require extraction of RNA but otherwise are analogous to DNA dot blot and Southern blot hybridization. Use of PCR for detection of c-*erb*B-2 mRNA has been described in two recent abstracts.[89,102]

Overproduction of c-*erb*B-2 mRNA can be measured by in situ hybridization. Sections are mounted on glass slides, treated with protease, hybridized with a radiolabeled probe, washed, treated with nuclease to remove unbound probe, and developed for autoradiography. Silver grains are seen only over tumor cells that overproduce c-*erb*B-2 mRNA. Negative control probes are used.[65,96,106] Our experience indicates that these techniques are relatively insensitive for detecting c-*erb*B-2 mRNA overproduction in routinely processed tissue. Although the sensitivity may be increased by modifications that allow simultaneous detection of c-*erb*B-2 DNA and mRNA, in situ hybridization still is cumbersome and expensive (unpublished data).

All of the above c-*erb*B-2 mRNA detection techniques have several problems that make them more difficult to perform than techniques for detecting DNA amplification. One major problem is the rapid degradation of RNA in tissue that is not immediately frozen or fixed. In addition, during the detection procedure, RNA can be degraded by RNase, a ubiquitous enzyme, which must be eliminated meticulously from laboratory solutions. Third, control probes to genes that are uniformly expressed in the tissue of interest need to be carefully selected.

## Detection of c-*erb*B-2 Protein Overproduction

The most accurate methods for detecting c-*erb*B-2 protein overproduction are the Western blot method and immunoprecipitation. Both techniques can document the binding specificity of various antibodies against c-*erb*B-2 protein. In Western blot studies, protein is extracted from the tissue, separated by electrophoresis (according to size), transferred to a membrane, and detected by using antibodies to c-*erb*B-2. In immunoprecipitation studies, antibodies against c-*erb*B-2 are added to a tumor lysate, and the resulting protein–antibody precipitate is separated by gel electrophoresis and stained for protein. Both Western blot and immunoprecipitation are useful research tools but currently are not practical for diagnostic pathology. Two recent abstracts have described an enzyme-linked immunosorbent assay (ELISA) for detection of c-*erb*B-2 protein.[19,45]

Overproduction of c-*erb*B-2 protein is most commonly assessed by various immunohistochemical techniques. These procedures often generate conflicting results, which are explained at least partially by three factors. First, various studies have used different polyclonal and monoclonal antibodies. Because some polyclonal antibodies recognize weak bands in addition to the c-*erb*B-2 protein band on Western blot or immunoprecipitation, the results of these studies should be interpreted with caution.[22,36,47,61] Even some monoclonal antibodies immunoprecipitate protein bands in addition to c-*erb*B-2 (p185).[30,59,86] Second, tissue fixation contributes to variability between studies. For example, some antibodies detect c-*erb*B-2 protein only in frozen tissue and do not react in fixed tissue. In general, formalin fixation diminishes the sensitivity of immunohistochemical methods and decreases the number of reactive cells.[61,86] When Bouin's fixative is used, there may be a higher percentage of positive cases.[22] Third, minimal criteria for interpreting immunohistochemical staining are generally lacking. Although there is general agreement that distinct crisp cytoplasmic membrane staining is diagnostic for c-*erb*B-2 activation in breast carcinoma, the number of positive cells and the staining intensity required to diagnose c-*erb*B-2 protein overproduction varies from study to study and from antibody to antibody. Degradation of c-*erb*B-2 protein is not a problem because it can be detected in intact form more than 24 hours after tumor resection without fixation or freezing.[64]

## ACTIVATION OF c-*erb*B-2 IN BREAST LESIONS

### Incidence of c-*erb*B-2 Activation

Most studies of c-*erb*B-2 oncogene activation do not specify histological subtypes of infiltrating breast carcinoma. Amplification of c-*erb*B-2 DNA was found in 19.1 percent (519 of 2715) of invasive carcinomas in 25 studies (Table 1), and c-*erb*B-2 mRNA or protein overproduction was detected in 20.9 percent (566 of 2714) of invasive carcinomas in 20 studies. Twelve studies have documented c-*erb*B-2 mRNA or protein overproduction in 15 percent (88 of 604) of carcinomas that lacked c-*erb*B-2 DNA amplification.

The incidence of c-*erb*B-2 activation in infiltrating breast carcinoma varies with the histological subtype. Approximately 22 percent (142 of 650) of infiltrating ductal carcinomas have c-*erb*B-2 activation, as expected from the above data. Other variants of breast carcinoma with frequent c-*erb*B-2 activation are inflammatory carcinoma (62 percent, 54 of 87), Paget's disease (82 percent, 9 of 11), and medullary carcinoma (22 percent, 5 of 23). In contrast, c-*erb*B-2 activation is infrequent in infiltrating lobular carcinoma (7 percent, 5 of 73) and tubular carcinoma (7 percent, 1 of 15).

The c-*erb*B-2 protein overproduction is present in 44 percent (44 of 100) of ductal carcinomas in situ and especially comedocarcinoma in situ (68 percent, 49 of 72). The micropapillary type of ductal carcinoma in situ also tends to have c-*erb*B-2 activation,[40,54,68] especially if larger cells are present. The greater fre-

**TABLE 1. *c-erbB-2* ACTIVATION IN MALIGNANT HUMAN BREAST NEOPLASMS**

| Histological Diagnosis | *c-erbB-2* DNA Amplification[a] | *c-erbB-2* mRNA Overproduction | *c-erbB-2* Protein Overproduction[b] |
|---|---|---|---|
| Carcinoma, not otherwise specified | 146/526,[81] 52/310,[17] 52/291,[109] 28/176,[67] 17/157,[113] 22/141,[35] 14/136,[57] 12/122,[4] 19/103,[79] 15/95,[90] 15/86,[111] 17/73,[77] 16/66,[42] 6/61,[50] 11/57,[52] 10/57,[65] 13/51,[13] 8/49,[21] 10/38,[92] 12/36,[94] 1/25,[15] 7/24,[91] 7/15,[31] 7/10,[96] 2/10[107] | 42/180,[69] 49/126,[36] 19/62,[65] 19/57,[50] 3/11,[90] 6/10,[96] 3/9[31] | 118/728,[85b] 58/330,[17b] 47/313,[86] 17/195,[11] 32/191,[58] 31/185,[101] 34/102,[42] 24/53,[52b] 23/47,[13] 22/45,[8] 11/36,[94] 7/24,[91] 1/10[61] |
| Carcinoma, type not specified but lacking *c-erbB-2* DNA amplification | — | 18/136,[81] 14/73,[34] 8/16,[65] 0/8,[90] 1/4,[31] 0/3[96] | 16/231,[17b] 18/136,[81] 13/35,[13] 14/29,[52b] 1/28,[92] 3/24,[94] 0/17[91] |
| Infiltrating ductal carcinoma | 21/118,[92] 23/107,[34] 17/50,[44] 7/37[93] 14/53 (comedo-carcinoma)[18] 3/33 (tubuloductal carcinoma)[16] | 35/95[34] | 22/137,[40] 14/93,[39] 9/34[88] |

| | | | |
|---|---|---|---|
| Inflammatory carcinoma | 33/80,[35] 3/6[52] | 46/75[35] | 5/6[52b] |
| Paget's disease | — | — | 5/6,[40] 2/3,[54] 2/2[92] |
| Tubular carcinoma | 0/5,[16] 0/1[93] | — | 1/9[40] |
| Medullary carcinoma | 2/4,[16] 0/1[34] | 0/1[34] | 1/12,[40] 1/3,[58] 1/2,[92] 0/1[39] |
| Mucinous carcinoma | 0/1,[16] 0/1[93] | — | 1/2[58] |
| Invasive papillary carcinoma | 0/2[93] | — | — |
| Infiltrating lobular carcinoma | 1/15,[16] 0/6[34] | 1/5[34] | 2/27,[92] 0/12,[40] 0/9,[39] 1/5[58] |
| Mammary fibrosarcoma | 0/1[93] | — | — |
| "Benign cystosarcoma" | — | — | 0/1[88] |
| Ductal CIS[c] with minimal invasion | 3/5[92] | — | — |
| Ductal CIS | 0/2[34] | 1/2[34] | 33/74,[40] 10/24[39] |
| Ductal CIS, solid or comedo type | — | — | 20/33,[58] 19/29,[92] 10/10[54] |
| Ductal CIS, micropapillary type | — | — | 10/10[58] |
| Ductal CIS, micropapillary or cribriform type | — | — | 1(focal)/14[54] |
| Ductal CIS, papillary or cribriform type | — | — | 0/16,[92] 1/9,[58] 0/3[40] |
| Lobular CIS | — | — | 0/16[40] |

[a]Shown as number of cases with activation/number of cases studied; reference is given as a superscript.
[b]These protein studies used Western blots; the rest used immunohistochemical methods.
[c]CIS = carcinoma in situ.

quency of c-*erb*B-2 protein overproduction in comedocarcinoma in situ, compared with infiltrating ductal carcinoma, could be explained by the fact that many infiltrating ductal carcinomas arise from other types of intraductal carcinoma, which show c-*erb*B-2 activation infrequently. Others have speculated that carcinoma in situ with c-*erb*B-2 activation tends to regress or to lose c-*erb*B-2 activation during progression to invasion.[40,68,92] Infiltrating and in situ components of ductal carcinoma, however, usually are similar with respect to c-*erb*B-2 activation,[11,39] although some authors have noted more heterogeneity of the immunohistochemical staining pattern in invasive than in in situ carcinoma.[40,42,88] Activation of c-*erb*B-2 is infrequent in lobular carcinoma in situ. If lesions contain more than one histological pattern of carcinoma in situ, the c-*erb*B-2 protein overproduction tends to occur in the comedocarcinoma in situ but may include other areas of carcinoma in situ.[42,54,68] Overproduction of c-*erb*B-2 protein in ductal carcinoma in situ correlates with larger cell size and a periductal lymphoid infiltrate.[68]

Activation of c-*erb*B-2 has not been identified in benign breast lesions, including fibrocystic disease, fibroadenomas, and radial scars (Table 2). Strong membrane immunohistochemical reactivity for c-*erb*B-2 has not been described in atypical ductal hyperplasia, although weak accentuation of membrane staining has been noted infrequently.[39,42,54] In normal breast tissue, c-*erb*B-2 DNA is diploid, and c-*erb*B-2 is expressed at lower levels than in activated tumors.[34,35,65,88]

These preliminary data suggest that c-*erb*B-2 activation may not be useful for resolving many of the common problems in diagnostic surgical pathology. For example, c-*erb*B-2 activation is infrequent in tubular carcinoma and radial scars. In addition, because c-*erb*B-2 activation is unusual in atypical ductal hyperplasia, cribriform carcinoma in situ, and papillary carcinoma in situ, detection of c-*erb*B-2 activation in these lesions may not be helpful in their differential diagnosis. The histological features of comedocarcinoma in situ, which commonly overproduces c-*erb*B-2, are unlikely to be mistaken for those of benign lesions. Activation of

**TABLE 2. c-*erb*B-2 ACTIVATION IN BENIGN HUMAN BREAST LESIONS**

| Histological Diagnosis | c-*erb*B-2 DNA Amplification[a] | c-*erb*B-2 mRNA Overproduction | c-*erb*B-2 Protein Overproduction |
|---|---|---|---|
| Fibrocystic disease | 0/10[93] | — | 0/32,[39] 0/9,[88] 0/8[65] |
| Atypical ductal hyperplasia | — | — | 2(weak)/21,[54] 1(cytoplasmic)/13[39] |
| Benign ductal hyperplasia | — | — | 0/12[39] |
| Sclerosing adenosis | — | — | 0/4[39] |
| Fibroadenomas | 0/16,[34] 0/6,[93] 0/2,[21] 0/1[91] | 0/6,[35] 0/3[34] | 0/21,[39] 0/10,[65] 0/8,[88] 0/3[42] |
| Radial scars | — | — | 0/22[39] |
| Blunt duct adenosis | — | — | 0/14[39] |
| "Breast mastosis" | — | 0/3[35] | — |

[a]Shown as number of cases with activation/number of cases studied; reference is given as a superscript.

c-*erb*B-2, however, does favor infiltrating ductal carcinoma over infiltrating lobular carcinoma. Further studies of these issues would be useful.

### Correlation of c-*erb*B-2 Activation With Pathologic Prognostic Factors

Multiple studies have attempted to correlate c-*erb*B-2 activation with various pathologic prognostic factors (Table 3). Activation of c-*erb*B-2 was correlated with lymph node metastasis in 8 of 28 series, with higher histological grade in 6 of 17 series, and with higher stage in 4 of 14 series. Large tumor size was not associated with c-*erb*B-2 activation in most studies (11 of 14). Tetraploid DNA content and low proliferation, measured by Ki-67, have been suggested as prognostic factors and may correlate with c-*erb*B-2 activation.[6,7]

### Correlation of c-*erb*B-2 Activation With Clinical Prognostic Factors

Various studies have attempted also to correlate c-*erb*B-2 activation with clinical features that may predict a poor outcome (Table 4). Activation of c-*erb*B-2 correlated with absence of estrogen receptors in 10 of 28 series and with absence of progesterone receptors in 6 of 18 series. In most studies, patient age did not correlate with c-*erb*B-2 activation, and, in the rest of the reports, c-*erb*B-2 activation was associated with either younger or older ages.

### Correlation of c-*erb*B-2 Activation With Patient Outcome

Slamon et al[79,81] first showed that amplification of the c-*erb*B-2 oncogene independently predicts decreased survival of patients with breast carcinoma. The correlation of c-*erb*B-2 amplification with poor outcome was nearly as strong as the correlation of number of involved lymph nodes with poor outcome. Slamon et al also reported that c-*erb*B-2 amplification is an important prognostic indicator only in patients with lymph node metastasis.[79,81]

A large number of subsequent studies also attempted to correlate c-*erb*B-2 activation with prognosis (Table 5). In 12 series, there was a correlation between c-*erb*B-2 activation and tumor recurrence or decreased survival. In five of these series, the predictive value of c-*erb*B-2 activation was reported to be independent of other prognostic factors. In contrast, 18 series did not confirm the correlation of c-*erb*B-2 activation with recurrence or survival. Four possible explanations for this controversy are discussed below.

One problem is that c-*erb*B-2 amplification correlates with prognosis mainly in patients with lymph node metastasis. As summarized in Table 5, most studies of patients with axillary lymph node metastasis showed a correlation of c-*erb*B-2 activation with poor outcome. In contrast, most studies of patients without axillary metastasis have not demonstrated a correlation with patient outcome. Table 6 summarizes the studies in which all patients (with and without axillary metastasis) were considered as one group. There is a trend for studies with a higher percentage of metastatic cases to show an association between c-*erb*B-2 activation and poor outcome. Thus, most of the current evidence suggests that c-*erb*B-2 activation has prognostic value only in patients with metastasis to lymph nodes.

**TABLE 3. CORRELATION OF c-*erb*B-2 ACTIVATION WITH PATHOLOGIC PROGNOSTIC FACTORS IN BREAST CARCINOMA**

| Prognostic Factor | P[a] | c-*erb*B-2 DNA Amplification[b] | c-*erb*B-2 mRNA Overproduction | c-*erb*B-2 Protein Overproduction[c] |
|---|---|---|---|---|
| Metastasis to axillary lymph nodes | <0.05 | (118)[35] (105)[34] (49)[21] | (104)[35] (92)[34] (9)[31] | (350)[85c] (36)[13] |
| | 0.05–0.15 | (103)[79] (86)[79] (58)[111] | — | (189)[92] |
| | >0.15 | (279)[17] (176)[87] (157)[113] (122)[4] (85)[90] (50)[52] (50)[44] (47)[15] (41)[93] | (50)[50] | (329)[17c] (261)[88] (195)[11] (185)[101] (102)[39] (50)[52b] |
| Larger size | <0.05 | (280)[17] | — | (330)[17c] (189)[92] |
| | 0.05–0.15 | (86)[79] | — | — |
| | >0.15 | (176)[87] (157)[113] (103)[79] (64)[77] (58)[111] (45)[21] | (51)[50] | (350)[85c] (185)[101] (34)[92] |
| Higher stage | <0.05 | (300)[17] (64)[77] (58)[111] | — | (349)[17c] |
| | 0.05–0.15 | (56)[52] | — | — |
| | >0.15 | (176)[87] (157)[113] (84)[90] (61)[50] (53)[21] (52)[57] (41)[93] | — | (102)[39] (56)[52b] |
| Higher histological grade | <0.05 | (47)[13] (15)[31] | (53)[65] | (176)[101] (168)[11] (38)[13] |
| | 0.05–0.15 | — | — | — |
| | >0.15 | (122)[4] (113)[34] (95)[90] (58)[111] (50)[44] (41)[93] | (86)[35] (65)[35] | (290)[85] (189)[92] (102)[39] |

[a]A correlation is statistically significant at <0.05, equivocal at best between 0.05 and 0.15, and not statistically significant at >0.15.
[b]Numbers inside parentheses are the number of patients in an individual study; superscript is the reference. Some studies analyzed more than one group of patients.
[c]By Western blot method; all other protein studies used immunohistochemical methods.

**TABLE 4. CORRELATION OF c-*erb*B-2 ACTIVATION WITH CLINICAL PROGNOSTIC FACTORS IN BREAST CARCINOMA**

| Prognostic Factor | P[a] | c-*erb*B-2 DNA Amplification[b] | c-*erb*B-2 mRNA Overproduction | c-*erb*B-2 Protein Overproduction[c] |
|---|---|---|---|---|
| Absence of estrogen receptors | <0.05 | (253)[109] (141)[35] (109)[34] (86)[79] (50)[44] (47)[13] | (104)[35] | (350)[85c] (330)[17a] (185)[101] |
| | 0.05–0.15 | — | — | — |
| | >0.15 | (157)[113] (122)[4] (103)[79] (95)[90] (64)[77] (61)[50] (58)[111] (53)[21] (51)[52] (41)[93] | (180)[69] (62)[65] (62)[85] (57)[50] | (290)[86] (172)[11] (51)[62c] (38)[13] |
| Absence of progesterone receptors | <0.05 | (253)[109] (141)[35] (109)[34] (50)[44] | — | (350)[85c] (306)[17a] |
| | 0.05–0.15 | (86)[79] (49)[52] | — | — |
| | >0.15 | (157)[113] (122)[4] (103)[79] (64)[77] | (180)[69] (103)[35] (62)[65] (56)[35] | (90)[11] (49)[52c] |
| Age (menopausal status) | <0.05 | — | — | (younger: 330)[17c] (older: 56)[52c] |
| | 0.05–0.15 | (younger: 86)[79] | — | — |
| | >0.15 | (230)[17] (176)[87] (157)[113] (122)[4] (116)[34] (103)[79] (95)[90] (64)[77] (58)[111] (56)[52] (53)[21] (49)[13] (41)[93] (15)[31] | (62)[65] | (350)[85c] (290)[86] (189)[92] (162)[11] (45)[92] |

[a]A correlation is statistically significant at <0.05, equivocal at best between 0.05 and 0.15, and not statistically significant at >0.15

[b]Numbers inside parentheses are the number of patients in an individual study; superscript is the reference. Some studies analyzed more than one group of patients.

[c]By Western blot method; all other protein studies used immunohistochemical methods.

**TABLE 5. CORRELATION OF c-*erb*B-2 ACTIVATION WITH OUTCOME IN PATIENTS WITH BREAST CARCINOMA**

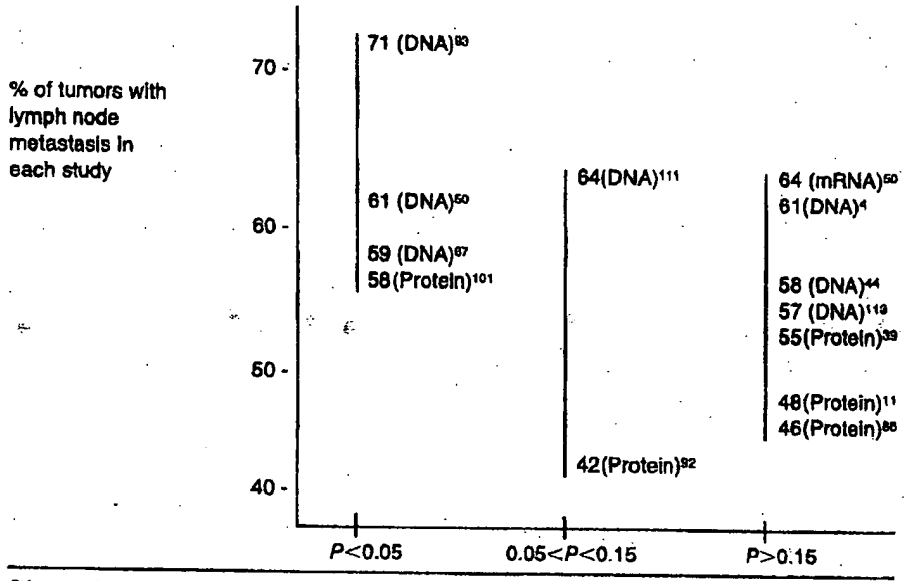
| | | Number of Patients | | | | |
|---|---|---|---|---|---|---|
| *P*[a] | Type of c-*erb*B-2 Activation[b] | *Total* | *With Metastastis to Axillary Lymph Nodes* | *No Metastasis* | Statistical Analysis[c] | Reference |
| <0.05 | DNA | 176 | | | M | 87 |
| <0.05 | DNA | 61 | | | U | 50 |
| <0.05 | DNA | 57 | | | U | 65 |
| <0.05 | DNA | 41 | | | U | 93 |
| <0.05 | mRNA | 62 | | | U | 65 |
| <0.05 | Protein | 102 | | | M | 101 |
| <0.05 | DNA | | 345 | | M | 81 |
| <0.05 | DNA | | 120 | | U | 17 |
| <0.05 | DNA | | 91 | | U | 87 |
| <0.05 | DNA | | 86 | | M | 79 |
| <0.05 | Protein-WB | | 350 | | M | 85 |
| <0.05 | Protein | | 62 | 44 | U | 101 |
| 0.05–0.15 | DNA | 57 | | | U | 111 |
| 0.05–0.15 | Protein | 189 | | | M | 92 |
| 0.05–0.15 | Protein | | 120 | | U | 86 |
| >0.15 | DNA | 130 | | | U | 113 |
| >0.15 | DNA | 122 | | | M | 4 |
| >0.15 | DNA | 50 | | | U | 44 |
| >0.15 | mRNA | 57 | | | U | 50 |
| >0.15 | Protein | 290 | | | M | 86 |
| >0.15 | Protein | 195 | | | U | 11 |
| >0.15 | Protein | 102 | | | U | 39 |
| >0.15 | Protein | | 137 | | U | 17 |
| >0.15 | DNA | | | 181 | M | 81 |
| >0.15 | DNA | | | 159 | U | 17 |
| >0.15 | DNA | | | 73 | U | 87 |
| >0.15 | Protein-WB | | | 378 | U | 85 |
| >0.15 | Protein-WB | | | 192 | U | 17 |
| >0.15 | Protein | | | 141 | U | 86 |
| >0.15 | Protein | | | 41 | U | 40 |

[a]The endpoints of these studies were tumor recurrence or decreased survival or both. Correlation between c-*erb*B-2 activation and a poorer patient outcome is statistically significant at <0.05, is of equivocal significance at 0.05 to 0.15, and is not significant at >0.15.

[b]Shown as variable measured. Letters "WB" indicate assay by Western blot; the other protein studies used immunohistochemical methods.

[c]M = multivariate statistical analysis; U = univariate statistical analysis.

**TABLE 6. PERCENTAGE OF BREAST CARCINOMAS WITH METASTASIS COMPARED WITH PROGNOSTIC SIGNIFICANCE OF c-*erb*B-2 ACTIVATION**




*P* for correlation of c-*erb*B-2 activation with patient outcome.
Each study's percentage of breast carcinomas with metastasis is compared with the correlation between c-*erb*B-2 activation and outcome. These data include only those studies that considered, as one group, all breast cancer patients, whether or not they had axillary metastasis. Superscripts are the references. In parentheses are the types of c-*erb*B-2 activation. *P* values are interpreted as in Table 3.

A second problem is that various types of breast carcinoma are grouped together in many survival studies. Because the current literature suggests that c-*erb*B-2 activation is infrequent in lobular carcinoma, studies that combine infiltrating ductal and lobular carcinomas may dilute the prognostic effect of c-*erb*B-2 activation in ductal tumors. In addition, most studies do not analyze inflammatory breast carcinoma separately. This condition frequently shows c-*erb*B-2 activation and has a worse prognosis than the usual mammary carcinoma, but it is an uncommon lesion.

A third potential problem is the paucity of studies that attempt to correlate c-*erb*B-2 activation with clinical outcome in subsets of breast carcinoma without metastasis. Two recent abstracts reported that in patients without lymph node metastasis who had various risk factors for recurrence (such as large tumor size and absence of estrogen receptors), c-*erb*B-2 overexpression predicted early recurrence.[23,67] In patients with ductal carcinoma in situ, one small study found no association between tumor recurrence and c-*erb*B-2 activation.[40]

A fourth problem is the lack of data regarding whether the prognosis correlates better with c-*erb*B-2 DNA amplification or with mRNA or protein overproduction. Most studies that find a correlation between c-*erb*B-2 activa-

tion and poor patient outcome measure c-*erb*B-2 DNA amplification (Table 5), and breast carcinoma patients with greater amplification of c-*erb*B-2 may have poorer survival.[79,81] Recent studies suggest that amplification has more prognostic power than overproduction,[17,34,35] but the clinical significance of c-*erb*B-2 overproduction without DNA amplification deserves further research.[17,52] Few studies have attempted to correlate patient outcome with c-*erb*B-2 mRNA overproduction, and many studies of c-*erb*B-2 protein overproduction use relatively less reliable methods such as immunohistochemical studies with polyclonal antibodies.

## Comparison of c-*erb*B-2 Activation With Other Oncogenes in Breast Carcinoma

Other oncogenes that may have prognostic implications in human breast cancer are reviewed elsewhere.[71,106] This section will be restricted to a comparison between the clinical relevance of c-*erb*B-2 and these other oncogenes.

The c-*myc* gene is often activated in breast carcinomas, but c-*myc* activation generally has less prognostic importance than c-*erb*B-2 activation.[21,34,77,87,93] One study found a correlation between increased mRNAs of c-*erb*B-2 and c-*myc*, although other reports have not confirmed this.[34,106] Subsequent research, however, could demonstrate a subset of breast carcinomas in which c-*myc* has more prognostic importance than c-*erb*B-2.

The gene c-*erb*B-1 for the epidermal growth factor receptor (EGFR) is homologous with c-*erb*B-2 but is infrequently amplified in breast carcinomas.[79] Overproduction of EGFR, however, occurs more frequently than amplification and may correlate with a poor prognosis. In studies that have examined both c-*erb*B-2 and EGFR in the same tumor, c-*erb*B-2 has a stronger correlation with poor prognostic factors.[35,52] Studies have tended to show no correlation between amplification of c-*erb*B-2 and c-*erb*B-1 or overproduction of c-*erb*B-2 and EGFR, although at the molecular level EGFR mediates phosphorylation of c-*erb*B-2 protein.[51,52,61,88,109] Recent reviews describe EGFR in breast carcinoma.[43,100]

The genes c-*erb*A and *ear*-1 are homologous to the thyroid hormone receptor, and they are located adjacent to c-*erb*B-2 on chromosome 17. These genes are frequently coamplified with c-*erb*B-2 in breast carcinomas. The absence of c-*erb*A expression in breast carcinomas, however, is evidence against an important role for this gene in breast neoplasia.[90] Amplification of c-*erb*B-2 can occur without *ear*-1 amplification, and these tumors have a decreased survival that is similar to tumors with both c-*erb*B-2 and *ear*-1 amplification.[87] Consequently, c-*erb*B-2 amplification seems to be more important than amplification of c-*erb*A or *ear*-1.

Other genes also have been compared with c-*erb*B-2 activation in breast carcinomas. One study found a significant correlation between increased c-*erb*B-2 mRNA and increased mRNAs of *fos*, platelet-derived growth factor chain A, and Ki-*ras*.[106] Allelic deletion of c-Ha-*ras* may indicate a poorer prognosis in breast carcinoma,[21] but it has not been compared with c-*erb*B-2 activation. Some studies have suggested a correlation between advanced stage or recurrence of breast carcinoma and activation of any one of several oncogenes.[21,113]

## ACTIVATION OF c-*erb*B-2 IN NON-MAMMARY TISSUES

### Incidence of c-*erb*B-2 Activation in Non-Mammary Tissues

Table 7 summarizes the normal tissues in which c-*erb*B-2 expression has been detected, usually with immunohistochemical methods using polyclonal anti-

**TABLE 7. PRESENCE OR ABSENCE OF c-*erb*B-2 mRNA OR c-*erb*B-2 PROTEIN IN NORMAL HUMAN TISSUES**

| Tissues With c-*erb*B-2 mRNA | Tissues Producing c-*erb*B-2 Protein[a] | Tissues Lacking c-*erb*B-2 mRNA | Tissues Lacking c-*erb*B-2 Protein |
|---|---|---|---|
| Skin[24] | Epidermis[56] | | |
| | External root sheath[55] | | |
| | Eccrine sweat gland[56] | | |
| | Fetal oral mucosa[62] | | Postnatal oral mucosa[62] |
| | Fetal esophagus[62] | | Postnatal esophagus[62] |
| Stomach[24] | Stomach[22,62] | | |
| | Fetal intestine[62a] | | |
| Jejunum[24] | Small intestine[22,62] | | |
| Colon[24] | Colon[22,62] | | |
| Kidney[24] | Fetal kidney[62a] | Kidneys[104] | Glomerulus[62] |
| | | | Postnatal Bowman's capsule[62] |
| | Fetal proximal tubule[62] | | Postnatal proximal tubule[62] |
| | Distal tubule[62] | | |
| | Fetal collecting duct[62] | | Postnatal collecting duct[62] |
| | Fetal renal pelvis[62] | | Postnatal renal pelvis[62] |
| | Fetal ureter[62] | | Postnatal fetal ureter[62] |
| Liver[24] | Hepatocytes[22] | | Liver[62,95] |
| | Pancreatic acini[22] | | |
| | Pancreatic ducts[22,62] | | |
| | Endocrine cells of islets of Langerhans[22] | | Pancreatic islets[62] |
| Lung[24] | Fetal trachea[62] | | Postnatal trachea[62] |
| | Fetal bronchioles[62] | | Postnatal bronchioles[62] |
| | Bronchioles[99] | | |
| | | | Postnatal alveoli[62,99] |
| Fetal brain[24] | | | Postnatal brain[62] |
| | Fetal ganglion cells[62] | | Postnatal ganglion cells[62] |
| Thyroid[1] | | | |
| Uterus[24] | | | |
| | Ovary[12] | | |
| | Blood vessels[42] | | Endothelium[62] |
| Placenta[24] | | | |
| | | | Adrenocortical cells[62] |
| | | | Postnatal thymus[62] |
| | | | Fibroblasts[62] |
| | | | Smooth muscle cells[62] |
| | | | Cardiac muscle cells[62] |

[a]This protein study used Western blots; the rest used immunohistochemical methods.

bodies. Only a few studies have been performed, and some of these do not demonstrate convincing cell membrane reactivity in the published photographs. The interpretations in these studies, however, are listed, with the caveat that these findings should be confirmed by immunoprecipitation or Western or RNA blots. Production of c-*erb*B-2 has been identified in normal epithelium of the gastrointestinal tract and skin. Discrepancies regarding c-*erb*B-2 protein in other tissues could be due, at least in part, to differences in techniques.

The data on c-*erb*B-2 activation in various non-mammary neoplasms should be interpreted with caution, because only small numbers of tumors have been studied, usually by immunohistochemical methods using polyclonal antibodies. Studies using cell lines have been excluded, because cell culture can induce amplification and overexpression of other genes, although this has not been documented for c-*erb*B-2.

Activation of c-*erb*B-2 has been identified in 32 percent (64 of 203) of ovarian carcinomas in eight studies (Table 8). One abstract[45] stated that ovarian carcinomas contained significantly more c-*erb*B-2 protein than ovarian non-epithelial malignancies. Another report[81] showed that 12 percent of ovarian carcinomas had c-*erb*B-2 overproduction without amplification.

Activation of c-*erb*B-2 has been identified in 20 percent (40 of 198) of gastric adenocarcinomas in seven studies, including 33 percent (21 of 64) of

**TABLE 8. c-*erb*B-2 ACTIVATION IN HUMAN GYNECOLOGIC TUMORS[a]**

| Tumor Type | c-*erb*B-2 DNA Amplification | c-*erb*B-2 mRNA Over-production | c-*erb*B-2 Protein Over-production |
|---|---|---|---|
| Ovary—carcinoma, not otherwise specified | 31/120,[81] 1/11,[57] 0/5,[107] 0/5,[84] 0/3,[112] 0/2,[72] 0/1[110] | 23/67[81] | 23/73,[12] 36/72[81] |
| Ovary—serous (papillary) carcinoma | 2/7,[110] 1/7,[112] 0/6[72] | — | — |
| Ovary—endometrioid carcinoma | 0/3[110] | — | — |
| Ovary—mucinous carcinoma | 1/2,[110] 0/1[72] | — | — |
| Ovary—clear cell carcinoma | 0/2,[112] 0/1[72] | — | — |
| Ovary—mixed epithelial carcinoma | 0/2[72] | — | — |
| Ovary—endometrioid borderline tumor | 0/1[72] | — | — |
| Ovary—mucinous borderline tumor | 0/3[72] | — | — |
| Ovary—serous cystadenoma | 0/4[72] | — | — |
| Ovary—mucinous cystadenoma | 0/2[72] | — | — |
| Ovary—sclerosing stromal tumor | 0/1[72] | — | — |
| Ovary—fibrothecoma | 0/1[72] | — | — |
| Uterus—endometrial adenocarcinoma | 0/4,[84] 0/1[110] | — | — |

[a]Shown as number of cases with amplification (or overproduction)/total number of cases studied; reference is given as superscript. All protein studies used immunohistochemical methods.

intestinal or tubular subtypes and 9 percent (4 of 47) of diffuse or signet ring cell subtypes (Table 9). Activation of c-*erb*B-2 has been detected in 2 percent (6 of 281) of colorectal carcinomas, although an additional immunohistochemical study detected c-*erb*B-2 protein in seven of eight tissues fixed in Bouin's solution. One study found greater immunohistochemical reactivity for c-*erb*B-2 protein in colonic adenomatous polyps than in the adjacent normal epithelium, using Bouin's fixative. Lesions with anaplastic features and progression to invasive carcinoma tended to show decreased immunohistochemical reactivity for c-*erb*B-2 protein.[22] Hepatocellular carcinomas (12 of 14 cases) and cholangiocarcinomas (46 of 63 cases) reacted with antibodies against c-*erb*B-2 in one study, but some of these "positive" cases showed only diffuse cytoplasmic staining, which

**TABLE 9. c-*erb*B-2 ACTIVATION IN HUMAN GASTROINTESTINAL TUMORS[a]**

| Tumor Type | c-*erb*B-2 DNA Amplification | c-*erb*B-2 Protein Overproduction |
|---|---|---|
| Esophagus—squamous cell carcinoma | 0/1[107] | 0/1[81] |
| Stomach—carcinoma, poorly differentiated | 0/22[108] | — |
| Stomach—adenocarcinoma | 2/24,[84] 2/9,[107] 2/8,[111] 2/8,[57] 0/1[108] | 4/27,[29] 3/10[61] |
| Stomach—carcinoma, intestinal or tubular type | 5/10[108] | 16/54[29] |
| Stomach—carcinoma, diffuse or signet ring cell type | 0/2[108] | 4/45[29] |
| Colorectum—carcinoma | 2/49,[84] 1/45,[111] 1/45,[57] 1/45,[60] 0/40,[81] 0/32,[107] 0/3[82] | 1/22,[58] 7/8[22b] |
| Colon—villous adenoma | 0/1[60] | — |
| Colon—tubulovillous adenoma | 0/5[60] | — |
| Colon—tubular adenoma | 0/7[60] | 19/19[22b] |
| Colon—hyperplastic polyp | 0/1[60] | — |
| Intestine—leiomyosarcoma | — | 0/1[61] |
| Hepatocellular carcinoma | 0/12[111] | 12/14,[85] 0/2[61] |
| Hepatoblastoma | 0/1[57] | — |
| Cholangiocarcinoma | — | 46/63[95] |
| Pancreas—adenocarcinoma | — | 2/80,[41c] 0/2[61] |
| Pancreas—acinar carcinoma | — | 0/1[41] |
| Pancreas—clear cell carcinoma | — | 0/2[41] |
| Pancreas—large cell carcinoma | — | 0/3[41] |
| Pancreas—signet ring carcinoma | — | 0/1[41] |
| Pancreas—chronic inflammation | — | 0/14[41c] |

[a]Shown as number of cases with amplification (or overproduction)/total number of cases studied; reference is given as superscript. All protein studies used immunohistochemical methods. No studies analyzed for c-*erb*B-2 mRNA.
[b]Tissues fixed in Bouin's solution.
[c]Only cases with distinct membrane staining are interpreted as showing c-*erb*B-2 overproduction.

**TABLE 10. c-*erb*B-2 ACTIVATION IN HUMAN PULMONARY TUMORS[a]**

| Tumor Type | c-*erb*B-2 DNA Amplification | c-*erb*B-2 Protein Overproduction |
|---|---|---|
| Non-small cell carcinoma | 2/60,[75] 0/60[81] | 1/84[59] |
| Epidermoid carcinoma | 0/13,[82] 0/10,[57] 0/6[20] | 3/5[99] |
| Adenocarcinoma | 0/21,[82] 1/13,[20] 0/7,[111] 0/7,[57] 0/3[107] | 4/12[99] |
| Large cell carcinoma | 0/9,[82] 0/6[20] | — |
| Small cell carcinoma | — | 0/26,[55] 0/3[99] |
| Carcinoid tumor | 0/1[82] | 0/3[99] |

[a]Shown as number of cases with amplification (or overproduction)/total number of cases studied; reference is given as superscript. All protein studies used immunohistochemical methods. No studies analyzed for c-*erb*B-2 mRNA.

does not indicate c-*erb*B-2 activation in breast neoplasms.[95] Also, some pancreatic carcinomas and chronic pancreatitis tissue had cytoplasmic immunohistochemical reactivity for c-*erb*B-2 protein, in addition to the rare case of pancreatic adenocarcinoma with distinct cell membrane staining.[41]

Tables 10 through 14 summarize the studies of c-*erb*B-2 activation in other neoplasms. The c-*erb*B-2 oncogene is not activated in most of these tumors. Activation of c-*erb*B-2 has been detected in 1 percent (4 of 299) of pulmonary non-small cell carcinomas in nine studies, although one additional report[99] found c-*erb*B-2 protein overproduction in 41 percent (7 of 17). Renal cell carcinoma had c-*erb*B-2 activation in 7 percent (2 of 30) in four studies. Overproduction of c-*erb*B-2 protein was described in one transitional cell carcinoma of the urinary bladder, a grade 2 papillary lesion.[58] Squamous cell carcinoma and basal cell carcinoma of the skin may contain c-*erb*B-2 protein, but it is not clear

**TABLE 11. c-*erb*B-2 ACTIVATION IN HUMAN HEMATOLOGIC PROLIFERATIONS[a]**

| Tumor Type | c-*erb*B-2 DNA Amplification | c-*erb*B-2 mRNA Over-production | c-*erb*B-2 Protein Over-production |
|---|---|---|---|
| Hematologic malignancies | 0/23[111] | — | — |
| Malignant lymphoma | 0/9,[57] 0/3[107] | 0/1[1] | 0/15[61] |
| Acute leukemia | 0/14[57] | — | — |
| Acute lymphoblastic leukemia | 0/1[107] | — | — |
| Acute myeloblastic leukemia | 0/3[107] | — | — |
| Chronic leukemia | 0/19[57] | — | — |
| Chronic lymphocytic leukemia | 0/6[107] | — | — |
| Chronic myelogenous leukemia | 0/8[107] | — | — |
| Myeloproliferative disorder | 0/1[57] | — | — |

[a]Shown as number of cases with amplification (or overproduction)/total number of cases studied; reference is given as superscript. All protein studies used immunohistochemical methods.

**TABLE 12. c-*erb*B-2 ACTIVATION IN HUMAN TUMORS OF SOFT TISSUE AND BONE[a]**

| Tumor Type | c-*erb*B-2 DNA Amplification |
|---|---|
| Sarcoma | 0/10,[111] 0/8[57] |
| Malignant fibrous histiocytoma | 0/1[107] |
| Liposarcoma | 0/3[107] |
| Pleomorphic sarcoma | 0/1[107] |
| Rhabdomyosarcoma | 0/1[107] |
| Osteogenic sarcoma | 0/2,[107] 0/2[57] |
| Chondrosarcoma | 0/1[107] |
| Ewing's sarcoma | 0/1[57] |
| Schwannoma | 0/1[57] |

[a]Shown as number of cases with amplification (or overproduction)/total number of cases studied; reference is given as superscript. No studies analyzed for c-*erb*B-2 mRNA or c-*erb*B-2 protein.

whether the protein level is increased over that of normal skin.[56] Thyroid carcinomas and adenomas can have low levels of increased c-*erb*B-2 mRNA. One abstract described low-level c-*erb*B-2 DNA amplification in one of ten salivary gland pleomorphic adenomas.[48]

## Correlation of c-*erb*B-2 Activation With Patient Outcome

Very few studies have attempted to correlate c-*erb*B-2 activation in non-mammary tumors with outcome. Slamon et al[81] showed that c-*erb*B-2 amplification or overexpression in ovarian carcinomas correlates with decreased survival, especially when marked activation is present. However, they did not report the stage, histological grade, or histological subtype of these neoplasms. Another study of stages III and IV ovarian carcinomas found a correlation between decreased survival and c-*erb*B-2 protein overproduction, but not between survival and histological grade.[12] One abstract stated that c-*erb*B-2 protein overproduction in 10 of 16 pulmonary adenocarcinomas correlated with decreased disease-free interval.[70] Another abstract described a tendency for immunohisto-

**TABLE 13. c-*erb*B-2 ACTIVATION IN HUMAN TUMORS OF THE URINARY TRACT[a]**

| Tumor Type | c-*erb*B-2 DNA Amplification | c-*erb*B-2 mRNA Overproduction | c-*erb*B-2 Protein Overproduction |
|---|---|---|---|
| Kidney—renal cell carcinoma | 1/5,[57] 1/4,[107] 0/5[84] | 0/16[104] | — |
| Wilms' tumor | 0/4[57] | — | — |
| Prostate—adenocarcinoma | — | — | 0/23[58] |
| Urinary bladder—carcinoma | — | — | 1/48[58] |

[a]Shown as number of cases with amplification (or overproduction)/total number of cases studied; reference is given as superscript. All protein studies used immunohistochemical methods.

TABLE 14. c-*erb*B-2 ACTIVATION IN MISCELLANEOUS HUMAN TUMORS[a]

| Tumor Type | c-*erb*B-2 DNA Amplification | c-*erb*B-2 mRNA Overproduction | c-*erb*B-2 Protein Overproduction |
|---|---|---|---|
| Skin—malignant melanoma | — | — | 0/10[58] |
| Skin, head and neck—squamous cell carcinoma | 0/7[107] | — | — |
| Site not stated—squamous cell carcinoma | 0/8,[57] 0/2[76] | — | — |
| Salivary gland—adenocarcinoma | 1/1[76] | — | — |
| Parotid gland—adenoid cystic carcinoma | — | — | 0/1[61] |
| Thyroid—anaplastic carcinoma | 0/1[1] | 0/1[1] | — |
| Thyroid—papillary carcinoma | 0/5[1] | 3(low levels)/5[1] | — |
| Thyroid—adenocarcinoma | 0/1[84] | — | — |
| Thyroid—adenoma | 0/2[1] | 1(low levels)/2[1] | — |
| Neuroblastoma | 0/35,[61] 0/9,[57] 0/1[76] | — | — |
| Meningioma | 0/2[57] | — | — |

[a]Shown as number of cases with amplification (or overproduction)/total number of cases studied; reference is given as superscript. All protein studies used immunohistochemical methods.

chemical reactivity for c-*erb*B-2 protein to correlate with higher grades of prostatic adenocarcinoma.[97] Additional prognostic studies of ovarian carcinomas and other neoplasms are needed.

## SUMMARY

Activation of the c-*erb*B-2 oncogene can occur by amplification of c-*erb*B-2 DNA and by overproduction of c-*erb*B-2 mRNA and c-*erb*B-2 protein. Approximately 20 percent of breast carcinomas show evidence of c-*erb*B-2 activation, which correlates with a poor prognosis primarily in patients with metastasis to axillary lymph nodes. Studies that have attempted to correlate c-*erb*B-2 activation with other prognostic factors in breast carcinoma have reported conflicting conclusions. The pathologic and clinical significance of c-*erb*B-2 activation in other neoplasms is unclear and should be assessed by additional studies.

## REFERENCES

1. Aasland R, Lillehaug JR, Male R, et al. Expression of oncogenes in thyroid tumors: Coexpression of c-*erb*B2/*neu* and c-*erb*B. *Br J Cancer.* 57:358, 1988
2. Akiyama T, Sudo C, Ogawara H, et al. The product of the human c-*erb*B-2 gene: A 185-kilodalton glycoprotein with tyrosine kinase activity. *Science.* 232:1644, 1986

3. Ali IU, Lidereau R, Theillet C, Callahan R. Reduction to homozygosity of genes on chromosome 11 in human breast neoplasia. *Science.* 238:185, 1987
4. Ali IU, Campbell G, Lidereau R, Callahan R. Amplification of c-*erb*B-2 and aggressive human breast tumors. *Science.* 240:1795, 1988
5. Ali IU, Campbell G, Lidereau R, Callahan R. Lack of evidence for the prognostic significance of c-*erb*B-2 amplification in human breast carcinoma. *Oncogene Res.* 3:139, 1988
6. Bacus SS, Bacus JW, Slamon DJ, Press MF. HER-2/*neu* oncogene expression and DNA ploidy analysis in breast cancer. *Arch Pathol Lab Med.* 114:164, 1990
7. Bacus SS, Ruby SG, Weinberg DS, et al. HER-2/*neu* oncogene expression and proliferation in breast cancers. *Am J Pathol.* 137:103, 1990
8. Bargmann CI, Hung MC, Weinberg RA. The *neu* oncogene encodes an epidermal growth factor receptor-related protein. *Nature.* 319:226, 1986
9. Bargmann CI, Hung MC, Weinberg RA. Multiple independent activations of the *neu* oncogene by a point mutation altering the transmembrane domain of p185. *Cell.* 45:649, 1986
10. Bargmann CI, Weinberg RA. Oncogenic activation of the *neu*-encoded receptor protein by point mutation and deletion. *EMBO J.* 7:2043, 1988
11. Barnes DM, Lammie GA, Millis RR, et al. An immunohistochemical evaluation of c-*erb*B-2 expression in human breast carcinoma. *Br J Cancer.* 58:448, 1988
12. Berchuck A, Kamel A, Whitaker R, et al. Overexpression of *HER-2/neu* is associated with poor survival in advanced epithelial ovarian cancer. *Cancer Res.* 50:4087, 1990
13. Berger MS, Locher GW, Saurer S, et al. Correlation of c-*erb*B-2 gene amplification and protein expression in human breast carcinoma with nodal status and nuclear grading. *Cancer Res.* 48:1238, 1988
14. Bernards R, Destree A, McKenzie S, et al. Effective tumor immunotherapy directed against an oncogene-encoded product using a vaccinia virus vector. *Proc Natl Acad Sci USA.* 84:6854, 1987
15. Biunno I, Pozzi MR, Pierotti MA, et al. Structure and expression of oncogenes in surgical specimens of human breast carcinomas. *Br J Cancer.* 57:464, 1988
16. Borg Å, Linell F, Idvall I, et al. *Her2/neu* amplification and comedo type breast carcinoma. *Lancet.* 1:1268, 1989
17. Borg Å, Tandon AK, Sigurdsson H, et al. *HER-2/neu* amplification predicts poor survival in node-positive breast cancer. *Cancer Res.* 50:4332, 1990
18. Bouchard L, Lamarre L, Tremblay PJ, Jolicoeur P. Stochastic appearance of mammary tumors in transgenic mice carrying the MMTV/c-*neu* oncogene. *Cell.* 57:931, 1989
19. Carney WP, Retos C, Petit D, et al. Quantitation of the *neu* oncogene protein using a monoclonal antibody based immunoassay (abstract). *Mod Pathol.* 3:15A, 1990
20. Cline MJ, Battifora H. Abnormalities of protooncogenes in non-small cell lung cancer: Correlations with tumor type and clinical characteristics. *Cancer.* 60:2669, 1987
21. Cline MJ, Battifora H, Yokota J. Proto-oncogene abnormalities in human breast cancer: Correlations with anatomic features and clinical course of disease. *J Clin Oncol.* 5:999, 1987
22. Cohen JA, Weiner DB, More KF, et al. Expression pattern of the *neu* (NGL) gene-encoded growth factor receptor protein ($p185^{neu}$) in normal and transformed epithelial tissues of the digestive tract. *Oncogene.* 4:81, 1989
23. Colnaghi MI, Miotti S, Andreola S, et al. New prognostic factors in breast cancer (abstract). *Am Assoc Cancer Res Ann Meeting.* 30:230A, 1989

24. Coussens L, Yang-Feng TL, Liao YC, et al. Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with *neu* oncogene. *Science.* 230:1132, 1985
25. Di Fiore PP, Pierce JH, Kraus MH, et al. *erb*B-2 is a potent oncogene when overexpressed in NIH/3T3 cells. *Science.* 237:178, 1987
26. Drebin JA, Link VC, Weinberg RA, Greene MI. Inhibition of tumor growth by a monoclonal antibody reactive with an oncogene-encoded tumor antigen. *Proc Natl Acad Sci USA.* 83:9129, 1986
27. Drebin JA, Link VC, Greene MI. Monoclonal antibodies reactive with distinct domains of the *neu* oncogene-encoded p185 molecule exert synergistic anti-tumor effects in vivo. *Oncogene.* 2:273, 1988
28. Drebin JA, Link VC, Greene MI. Monoclonal antibodies specific for the *neu* oncogene product directly mediate anti-tumor effects in vivo. *Oncogene.* 2:387, 1988
29. Falck VG, Gullick WJ. c-*erb*B-2 oncogene product staining in gastric adenocarcinoma. An immunohistochemical study. *J Pathol.* 159:107, 1989
30. Fendly BM, Winget M, Hudziak RM, et al. Characterization of murine monoclonal antibodies reactive to either the human epidermal growth factor receptor or *HER2/neu* gene product. *Cancer Res.* 50:1550, 1990
31. Fontaine J, Tesseraux M, Klein V, et al. Gene amplification and expression of the *neu* (c-*erb*B-2) sequence in human mammary carcinoma. *Oncology.* 45:360, 1988
32. Frye RA, Benz CC, Liu E. Detection of amplified oncogenes by differential polymerase chain reaction. *Oncogene.* 4:1153, 1989
33. Fukushige SI, Matsubara KI, Yoshida M, et al. Localization of a novel v-*erb*B-related gene, c-*erb*B-2, on human chromosome 17 and its amplification in a gastric cancer cell line. *Mol Cell Biol.* 6:955, 1986
34. Guerin M, Barrois M, Terrier MJ, et al. Overexpression of either c-*myc* or c-*erb*B-2/*neu* proto-oncogenes in human breast carcinomas: Correlation with poor prognosis. *Oncogene Res.* 3:21, 1988
35. Guerin M, Gabillot M, Mathieu MC, et al. Structure and expression of c-*erb*B-2 and EGF receptor genes in inflammatory and non-inflammatory breast cancer: Prognostic significance. *Int J Cancer.* 43:201, 1989
36. Gullick WJ, Berger MS, Bennett PLP, et al. Expression of the c-*erb*B-2 protein in normal and transformed cells. *Int J Cancer.* 40:246, 1987
37. Gullick WJ, Venter DJ. The c-*erb*B2 gene and its expression in human cancers. In: Waxman J, Sikora K, eds. *The Molecular Biology of Cancer.* Boston, Blackwell Sci Publ; 1989: 38–53
38. Gullick WJ. Expression of the c-*erb*B-2 proto-oncogene protein in human breast cancer. *Recent Results Cancer Res.* 113:51, 1989
39. Gusterson BA, Machin LG, Gullick WJ, et al. c-*erb*B-2 expression in benign and malignant breast disease. *Br J Cancer.* 58:453, 1988
40. Gusterson BA, Machin LG, Gullick WJ, et al. Immunohistochemical distribution of c-*erb*B-2 in infiltrating and in situ breast cancer. *Int J Cancer.* 42:842, 1988
41. Hall PA, Hughes CM, Staddon SL, et al. The c-*erb*B-2 proto-oncogene in human pancreatic cancer. *J Pathol.* 161:195, 1990
42. Hanna W, Kahn HJ, Andrulis I, Pawson T. Distribution and patterns of staining of *neu* oncogene product in benign and malignant breast diseases. *Mod Pathol.* 3:455, 1990
43. Harris AL, Nicholson S. Epidermal growth factor receptors in human breast cancer.

In: Lippman ME, Dickson RB, eds. *Breast Cancer: Cellular and Molecular Biology.* Boston, Kluwer Academic Publ; 1988: 93–118

44. Heintz NH, Leslie KO, Rogers LA, Howard PL. Amplification of the c-*erb*B-2 oncogene and prognosis of breast adenocarcinoma. *Arch Pathol Lab Med.* 114:160, 1990
45. Huettner P, Carney W, Delellis R, et al. Quantification of the *neu* oncogene product in ovarian neoplasms (abstract). *Mod Pathol.* 3:46A, 1990
46. Hung MC, Yan DH, Zhao X. Amplification of the proto-*neu* oncogene facilitates oncogenic activation by a single point mutation. *Proc Natl Acad Sci USA.* 86:2545, 1989
47. Hynes NE, Gerber HA, Saurer S, Groner B. Overexpression of the c-*erb*B-2 protein in human breast tumor cell lines. *J Cell Biochem.* 39:167, 1989
48. Kahn HJ, Hanna W, Auger M, Andreulis I. Expression and amplification of *neu* oncogene in pleomorphic adenoma of salivary glands (abstract). *Mod Pathol.* 3:50A, 1990
49. King CR, Kraus MH, Aaronson SA. Amplification of a novel v-*erb*B-related gene in a human mammary carcinoma. *Science.* 229:974, 1985
50. King CR, Swain SM, Porter L, et al. Heterogeneous expression of *erb*B-2 messenger RNA in human breast cancer. *Cancer Res.* 49:4185, 1989
51. Kokai Y, Dobashi K, Weiner DB, et al. Phosphorylation process induced by epidermal growth factor receptor alters the oncogenic and cellular *neu* (*NGL*) gene products. *Proc Natl Acad Sci USA.* 85:5389, 1988
52. Lacroix H, Iglehart JD, Skinner MA, Kraus MH. Overexpression of *erb*B-2 or EGF receptor proteins present in early stage mammary carcinoma is detected simultaneously in matched primary tumors and regional metastases. *Oncogene.* 4:145, 1989
53. Lemoine NR, Staddon S, Dickson C, et al. Absence of activating transmembrane mutations in the c-*erb*B-2 proto-oncogene in human breast cancer. *Oncogene.* 5:237, 1990
54. Lodato RF, Maguire HC, Greene MJ, et al. Immunohistochemical evaluation of c-*erb*B-2 oncogene expression in ductal carcinoma in situ and atypical ductal hyperplasia of the breast. *Mod Pathol.* 3:449, 1990
55. Maguire HC, Greene MI. The *neu* (c-*erb*B-2) oncogene. *Semin Oncol.* 16:148, 1989
56. Maguire HC, Jaworsky C, Cohen JA, et al. Distribution of *neu* (c-*erb*B-2) protein in human skin. *J Invest Dermatol.* 89:786, 1989
57. Masuda H, Battifora H, Yokota J, et al. Specificity of proto-oncogene amplification in human malignant diseases. *Mol Biol Med.* 4:213, 1987
58. McCann A, Dervan PA, Johnston PA, et al. c-*erb*B-2 oncoprotein expression in primary human tumors. *Cancer.* 65:88, 1990
59. McKenzie SJ, Marks PJ, Lam T, et al. Generation and characterization of monoclonal antibodies specific for the human *neu* oncogene product, p185. *Oncogene.* 4:543, 1989
60. Meltzer SJ, Ahnen DJ, Battifora H, et al. Protooncogene abnormalities in colon cancers and adenomatous polyps. *Gastroenterology.* 92:1174, 1987
61. Mori S, Akiyama T, Morishita Y, et al. Light and electron microscopical demonstration of c-*erb*B-2 gene product-like immunoreactivity in human malignant tumors. *Virchows Arch* [*B*]. 54:8, 1987
62. Mori S, Akiyama T, Yamada Y, et al. C-*erb*B-2 gene product, a membrane protein commonly expressed in human fetal epithelial cells. *Lab Invest.* 61:93, 1989
63. Muller WJ, Sinn E, Pattengale PK, et al. Single-step induction of mammary

adenocarcinoma in transgenic mice bearing the activated c-*neu* oncogene. *Cell.* 54:105, 1988

64. Ong G, Gullick W, Sikora K. Oncoprotein stability after tumor resection. *Br J Cancer.* 61:538, 1990
65. Parks HC, Lillycrop K, Howell A, Craig RK. C-*erb*B2 mRNA expression in human breast tumors: Comparison with c-*erb*B2 DNA amplification and correlation with prognosis. *Br J Cancer.* 61:39, 1990
66. Popescu NC, King CR, Kraus MH. Localization of the *erb*B-2 gene on normal and rearranged chromosomes 17 to bands q12–21.32. *Genomics.* 4:362, 1989
67. Press MF, Pike MC, Paterson MC, et al. Overexpression of *HER-2/neu.* proto-oncogene in node negative breast cancer: Correlation with increased risk of early recurrent disease (abstract). *Mod Pathol.* 3:80A, 1990
68. Ramachandra S, Machin L, Ashley S, et al. Immunohistochemical distribution of c-*erb*B-2 in in situ breast carcinoma: A detailed morphological analysis. *J Pathol.* 161:7, 1990
69. Rio MC, Bellocq JP, Gairard B, et al. Specific expression of the pS2 gene in subclasses of breast cancers in comparison with expression of the estrogen and progesterone receptors and the oncogene *ERBB2. Proc Natl Acad Sci USA.* 84:9243, 1987
70. Robinson R, Kern J, Weiner D, et al. p185$^{neu}$ expression in human lung non-small cell carcinomas: An immunohistochemical study with clinicopathologic correlation (abstract). *Mod Pathol.* 3:85A, 1990
71. Rochlitz CF, Benz CC. Oncogenes in human solid tumors. In: Benz C, Liu E, eds. *Oncogenes.* Boston, Kluwer Academic Publ; 1989: 199–240
72. Sasano H, Garret CT, Wilkinson DS, et al. Protooncogene amplification and tumor ploidy in human ovarian neoplasms. *Hum Pathol.* 21:382, 1990
73. Schechter AL, Stern DF, Vaidyanathan L, et al. The *neu* oncogene: An *erb*-B-related gene encoding a 185,000-$M_r$ tumor antigen. *Nature.* 312:513, 1984
74. Schechter AL, Hung MC, Vaidyanathan L, et al. The *neu* gene: An *erb*B-homologous gene distinct from and unlinked to the gene encoding the EGF receptor. *Science.* 229:976, 1985
75. Schneider PM, Hung MC, Chiocca SM, et al. Differential expression of the c-*erb*B-2 gene in human small cell and non-small cell lung cancer. *Cancer Res.* 49:4968, 1989
76. Semba K, Kamata N, Toyoshima K, Yamamoto T. A v-*erb*B-related protooncogene, c-*erb*B-2, is distinct from the c-*erb*B-1/epidermal growth factor-receptor gene and is amplified in a human salivary gland adenocarcinoma. *Proc Natl Acad Sci USA.* 82:6497, 1985
77. Seshadri R, Matthews C, Dobrovic A, Horsfall DJ. The significance of oncogene amplification in primary breast cancer. *Int J Cancer.* 43:270, 1989
78. Shih C, Padhy LC, Murray M, Weinberg RA. Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts. *Nature.* 290:261, 1981
79. Slamon DJ, Clark GM, Wong SG, et al. Human breast cancer: Correlation of relapse and survival with amplification of the *HER-2/neu* oncogene. *Science.* 235:177, 1987
80. Slamon DJ, Clark GM. Amplification of c-*erb*B-2 and aggressive human breast tumors. *Science.* 240:1795, 1988
81. Slamon DJ, Godolphin W, Jones LA, et al. Studies of the *HER-2/neu* proto-oncogene in human breast and ovarian cancer. *Science.* 244:707, 1989

82. Slebos RJC, Evers SG, Wagenaar SS, Rodenhuis S. Cellular protooncogenes are infrequently amplified in untreated non-small cell lung cancer. *Br J Cancer.* 59:76, 1989
83. Stern DF, Kamps MP, Cao H. Oncogenic activation of p185$^{neu}$ stimulates tyrosine phosphorylation in vivo. *Mol Cell Biol.* 8:3969, 1988
84. Tal M, Wetzler M, Josefberg Z, et al. Sporadic amplification of the *HER2/neu* protooncogene in adenocarcinomas of various tissues. *Cancer Res.* 48:1517, 1988
85. Tandon AK, Clark GM, Chamness GC, et al. *HER-2/neu* oncogene protein and prognosis in breast cancer. *J Clin Oncol.* 7:1120, 1989
86. Thor AD, Schwartz LH, Koerner FC, et al. Analysis of c-*erb*B-2 expression in breast carcinomas with clinical follow-up. *Cancer Res.* 49:7147, 1989
87. Tsuda H, Hirohashi S, Shimosato Y, et al. Correlation between long-term survival in breast cancer patients and amplification of two putative oncogene-coamplification units: *hst-1/int-2* and c-*erb*B-2/*ear-1*. *Cancer Res.* 49:3104, 1989
88. Tsutsumi Y, Naber SP, DeLellis RA, et al. *Neu* oncogene protein and epidermal growth factor receptor are independently expressed in benign and malignant breast tissues. *Hum Pathol.* 21:750, 1990
89. Tsutsumi Y, Stork PJ, Wolfe HJ. Detection of DNA amplification and mRNA overexpression of the *neu* oncogene in breast carcinomas by polymerase chain reaction (abstract). *Mod Pathol.* 3:101A, 1990
90. Van de Vijver M, van de Bersselaar R, Devilee P, et al. Amplification of the *neu* (c-*erb*B-2) oncogene in human mammary tumors is relatively frequent and is often accompanied by amplification of the linked c-*erb*A oncogene. *Mol Cell Biol.* 7:2019, 1987
91. Van de Vijver MJ, Mooi WJ, Wisman P, et al. Immunohistochemical detection of the *neu* protein in tissue sections of human breast tumors with amplified *neu* DNA. *Oncogene.* 2:175, 1988
92. Van de Vijver MJ, Peterse JL, Mooi WJ, et al. *Neu*-protein overexpression in breast cancer: Association with comedo-type ductal carcinoma in situ and limited prognostic value in stage II breast cancer. *N Engl J Med.* 319:1239, 1988
93. Varley JM, Swallow JE, Brammar WJ, et al. Alterations to either c-*erb*B-2 (*neu*) or c-*myc* proto-oncogenes in breast carcinomas correlate with poor short-term prognosis. *Oncogene.* 1:423, 1987
94. Venter DJ, Tuzi NL, Kumar S, Gullick WJ. Overexpression of the c-*erb*B-2 oncoprotein in human breast carcinomas: Immunohistological assessment correlates with gene amplification. *Lancet.* 2:69, 1987
95. Voravud N, Foster CS, Gilbertson JA, et al. Oncogene expression in cholangiocarcinoma and in normal hepatic development. *Hum Pathol.* 20:1163, 1989
96. Walker RA, Senior PV, Jones JL, et al. An immunohistochemical and in situ hybridization study of *c-myc* and *c-erb*B-2 expression in primary human breast carcinomas. *J Pathol.* 158:97, 1989
97. Ware JL, Maygarden SJ, Koontz WW, Strom SC. Differential reactivity with anti-c-*erb*B-2 antiserum among human malignant and benign prostatic tissue (abstract). *Am Assoc Cancer Res Ann Meeting.* 30:437A, 1989
98. Weiner DB, Liu J, Cohen JA, et al. A point mutation in the *neu* oncogene mimics ligand induction of receptor aggregation. *Nature.* 339:230, 1989
99. Weiner DB, Nordberg J, Robinson R, et al. Expression of the *neu* gene-encoded protein (p185$^{neu}$) in human non-small cell carcinomas of the lung. *Cancer Res.* 50:421, 1990

100. Wells A. The epidermal growth factor receptor and its ligand. In: Benz C, Liu E, eds. *Oncogenes.* Boston, Kluwer Academic Pub; 1989: 143–168
101. Wright C, Angus B, Nicholson S, et al. Expression of c-*erb*B-2 oncoprotein: A prognostic indicator in human breast cancer. *Cancer Res.* 49:2087, 1989
102. Wu A, Colombero A, Low J, et al. Analysis of expression and mutation of the *erb*B-2 gene in breast carcinoma by the polymerase chain reaction (abstract). *Mod Pathol.* 3:108A, 1990
103. Yamamoto T, Ikawa S, Akiyama T, et al. Similarity of protein encoded by the human c-*erb*B-2 gene to epidermal growth factor receptor. *Nature.* 319:230, 1986
104. Yao M, Shuin T, Misaki H, Kubota Y. Enhanced expression of c-*myc* and epidermal growth factor receptor (C-*erb*B-1) genes in primary human renal cancer. *Cancer Res.* 48:6753, 1988
105. Yarden Y, Weinberg RA. Experimental approaches to hypothetical hormones: Detection of a candidate ligand of the *neu* protooncogene. *Proc Natl Acad Sci USA.* 86:3179, 1989
106. Yee LD, Kacinski BM, Carter D. Oncogene structure, function and expression in breast cancer. *Semin Diagn Pathol.* 6:110, 1989
107. Yokota J, Yamamoto T, Toyoshima K, et al. Amplification of c-*erb*B-2 oncogene in human adenocarcinomas in vivo. *Lancet.* 1:765, 1986
108. Yokota J, Yamamoto T, Miyajima N, et al. Genetic alterations of the c-*erb*B-2 oncogene occur frequently in tubular adenocarcinoma of the stomach and are often accompanied by amplification of the v-*erb*A homologue. *Oncogene.* 2:283, 1988
109. Zeillinger R, Kury F, Czerwenka K, et al. *HER*-2 amplification, steroid receptors and epidermal growth factor receptor in primary breast cancer. *Oncogene.* 4:109, 1989
110. Zhang X, Silva E, Gershenson D, Hung MC. Amplification and rearrangement of c-*erb*B proto-oncogenes in cancer of human female genital tract. *Oncogene.* 4:985, 1989
111. Zhou D, Battifora H, Yokota J, et al. Association of multiple copies of the c-*erb*B-2 oncogene with spread of breast cancer. *Cancer Res.* 47:6123, 1987
112. Zhou D, Gonzalez-Cadavid N, Ahuja H, et al. A unique pattern of proto-oncogene abnormalities in ovarian adenocarcinomas. *Cancer.* 62:1573, 1988
113. Zhou D, Ahuja H, Cline MJ. Proto-oncogene abnormalities in human breast cancer: c-*erb*B-2 amplification does not correlate with recurrence of disease. *Oncogene.* 4:105, 1989

8 8

GNE2930R1C3

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

COPY

| | | |
|---|---|---|
| Applicant | : | Botstein, et al. |
| Appl. No. | : | 10/032,996 |
| Filed | : | December 27, 2001 |
| For | : | SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME |
| Examiner | : | Fredman, J. |
| Group Art Unit | : | 1634 |

DECLARATION OF PAUL POLAKIS, PH.D. UNDER 37 C.F.R. § 1.132

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Attached is the Declaration of Paul Polakis, Ph.D.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: June 16, 2004

By: AnneMarie Kaiser

AnneMarie Kaiser
Registration No. 37,649
Attorney of Record
Customer No. 30,313
(619) 235-8550

S:\DOCS\BSG\BSG-1215.DOC
061404

# DECLARATION OF PAUL POLAKIS, Ph.D.

I, Paul Polakis, Ph.D., declare and say as follows:

1. I was awarded a Ph.D. by the Department of Biochemistry of the Michigan State University in 1984. My scientific Curriculum Vitae is attached to and forms part of this Declaration (Exhibit A).

2. I am currently employed by Genentech, Inc. where my job title is Staff Scientist. Since joining Genentech in 1999, one of my primary responsibilities has been leading Genentech's Tumor Antigen Project, which is a large research project with a primary focus on identifying tumor cell markers that find use as targets for both the diagnosis and treatment of cancer in humans.

3. As part of the Tumor Antigen Project, my laboratory has been analyzing differential expression of various genes in tumor cells relative to normal cells. The purpose of this research is to identify proteins that are abundantly expressed on certain tumor cells and that are either (i) not expressed, or (ii) expressed at lower levels, on corresponding normal cells. We call such differentially expressed proteins "tumor antigen proteins". When such a tumor antigen protein is identified, one can produce an antibody that recognizes and binds to that protein. Such an antibody finds use in the diagnosis of human cancer and may ultimately serve as an effective therapeutic in the treatment of human cancer.

4. In the course of the research conducted by Genentech's Tumor Antigen Project, we have employed a variety of scientific techniques for detecting and studying differential gene expression in human tumor cells relative to normal cells, at genomic DNA, mRNA and protein levels. An important example of one such technique is the well known and widely used technique of microarray analysis which has proven to be extremely useful for the identification of mRNA molecules that are differentially expressed in one tissue or cell type relative to another. In the course of our research using microarray analysis, we have identified approximately 200 gene transcripts that are present in human tumor cells at significantly higher levels than in corresponding normal human cells. To date, we have generated antibodies that bind to about 30 of the tumor antigen proteins expressed from these differentially expressed gene transcripts and have used these antibodies to quantitatively determine the level of production of these tumor antigen proteins in both human cancer cells and corresponding normal cells. We have then compared the levels of mRNA and protein in both the tumor and normal cells analyzed.

5. From the mRNA and protein expression analyses described in paragraph 4 above, we have observed that there is a strong correlation between changes in the level of mRNA present in any particular cell type and the level of protein

expressed from that mRNA in that cell type. In approximately 80% of our observations we have found that increases in the level of a particular mRNA correlates with changes in the level of protein expressed from that mRNA when human tumor cells are compared with their corresponding normal cells.

6. Based upon my own experience accumulated in more than 20 years of research, including the data discussed in paragraphs 4 and 5 above and my knowledge of the relevant scientific literature, it is my considered scientific opinion that for human genes, an increased level of mRNA in a tumor cell relative to a normal cell typically correlates to a similar increase in abundance of the encoded protein in the tumor cell relative to the normal cell. In fact, it remains a central dogma in molecular biology that increased mRNA levels are predictive of corresponding increased levels of the encoded protein. While there have been published reports of genes for which such a correlation does not exist, it is my opinion that such reports are exceptions to the commonly understood general rule that increased mRNA levels are predictive of corresponding increased levels of the encoded protein.

7. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful statements may jeopardize the validity of the application or any patent issued thereon.

Dated: 5/07/04

By: Paul Polakis

Paul Polakis, Ph.D.

SV 2031808 v1

# EXHIBIT A

# CURRICULUM VITAE

PAUL G. POLAKIS
Staff Scientist
Genentech, Inc
1 DNA Way, MS#40
S. San Francisco, CA 94080

## EDUCATION:

Ph.D., Biochemistry, Department of Biochemistry,
Michigan State University (1984)

B.S., Biology. College of Natural Science, Michigan State University (1977)

## PROFESSIONAL EXPERIENCE:

| | |
|---|---|
| 2002-present | Staff Scientist, Genentech, Inc<br>S. San Francisco, CA |
| 1999- 2002 | Senior Scientist, Genentech, Inc.,<br>S. San Francisco, CA |
| 1997 -1999 | Research Director<br>Onyx Pharmaceuticals, Richmond, CA |
| 1992- 1996 | Senior Scientist, Project Leader, Onyx<br>Pharmaceuticals, Richmond, CA |
| 1991-1992 | Senior Scientist, Chiron Corporation,<br>Emeryville, CA. |
| 1989-1991 | Scientist, Cetus Corporation, Emeryville CA. |
| 1987-1989 | Postdoctoral Research Associate, Genentech,<br>Inc., South SanFrancisco, CA. |
| 1985-1987 | Postdoctoral Research Associate, Department<br>of Medicine, Duke University Medical Center,<br>Durham, NC |

| | |
|---|---|
| 1984-1985 | Assistant Professor, Department of Chemistry, Oberlin College, Oberlin, Ohio |
| 1980-1984 | Graduate Research Assistant, Department of Biochemistry, Michigan State University East Lansing, Michigan |

**PUBLICATIONS:**

1. **Polakis, P G.** and Wilson, J. E. 1982 Purification of a Highly Bindable Rat Brain Hexokinase by High Performance Liquid Chromatography. **Biochem. Biophys. Res. Commun**. 107, 937-943.

2. **Polakis, P.G**. and Wilson, J. E. 1984 Proteolytic Dissection of Rat Brain Hexokinase: Determination of the Cleavage Pattern during Limited Digestion with Trypsin. **Arch. Biochem. Biophys**. 234, 341-352.

3. **Polakis, P. G.** and Wilson, J. E. 1985 An Intact Hydrophobic N-Terminal Sequence is Required for the Binding Rat Brain Hexokinase to Mitochondria. **Arch. Biochem. Biophys**. 236, 328-337.

4. Uhing, R.J., **Polakis,P.G**. and Snyderman, R. 1987 Isolaton of GTP-binding Proteins from Myeloid HL60 Cells. **J. Biol. Chem**. 262, 15575-15579.

5. **Polakis, P.G.**, Uhing, R.J. and Snyderman, R. 1988 The Formylpeptide Chemoattractant Receptor Copurifies with a GTP-binding Protein Containing a Distinct 40 kDa Pertussis Toxin Substrate. **J. Biol. Chem**. 263, 4969-4979.

6. Uhing, R. J., Dillon, S., **Polakis, P. G.**, Truett, A. P. and Snyderman, R. 1988 Chemoattractant Receptors and Signal Transduction Processes in Cellular and Molecular Aspects of Inflammation ( Poste, G. and Crooke, S. T. eds.) pp 335-379.

7. **Polakis, P.G.**, Evans, T. and Snyderman 1989 Multiple Chromatographic Forms of the Formylpeptide Chemoattractant Receptor and their Relationship to GTP-binding Proteins. **Biochem. Biophys. Res. Commun**. 161, 276-283.

8. **Polakis, P. G.**, Snyderman, R. and Evans, T. 1989 Characterization of G25K, a GTP-binding Protein Containing a Novel Putative Nucleotide Binding Domain. **Biochem. Biophys. Res. Comun.** 160, 25-32.

9. **Polakis, P.**, Weber,R.F., Nevins,B., Didsbury, J. Evans,T. and Snyderman, R. 1989 Identification of the ral and rac1 Gene Products, Low Molecular Mass GTP-binding Proteins from Human Platelets. **J. Biol. Chem**. 264, 16383-16389.

10. Snyderman, R., Perianin, A., Evans, T., **Polakis, P.** and Didsbury, J. 1989 G Proteins and Neutrophil Function. In ADP-Ribosylating Toxins and G Proteins: Insights into Signal Transduction. ( J. Moss and M. Vaughn, eds.) Amer. Soc. Microbiol. pp. 295-323.

**11.** Hart, M.J., **Polakis, P.G.**, Evans, T. and Cerrione, R.A. 1990 The Identification and Charaterization of an Epidermal Growth Factor-Stimulated Phosphorylation of a Specific Low Molecular Mass GTP-binding Protein in a Reconstituted Phospholipid Vesicle System. **J. Biol. Chem**. 265, 5990-6001.

**12.** Yatani, A., Okabe, K., **Polakis, P.** Halenbeck, R. McCormick, F. and Brown, A. M. 1990 ras p21 and GAP Inhibit Coupling of Muscarinic Receptors to Atrial $K^+$ Channels. **Cell**. 61, 769-776.

**13.** Munemitsu, S., Innis, M.A., Clark, R., McCormick, F., Ullrich, A. and **Polakis, P.G.** 1990 Molecular Cloning and Expression of a G25K cDNA, the Human Homolog of the Yeast Cell Cycle Gene CDC42. **Mol. Cell. Biol**. 10, 5977-5982.

**14. Polakis, P.G.** Rubinfeld, B. Evans, T. and McCormick, F. 1991 Purification of Plasma Membrane-Associated GTPase Activating Protein (GAP) Specific for rap-1/krev-1 from HL60 Cells. **Proc. Natl. Acad. Sci. USA** 88, 239-243.

**15.** Moran, M. F., **Polakis, P.**, McCormick, F., Pawson, T. and Ellis, C. 1991 Protein Tyrosine Kinases Regulate the Phosphorylation, Protein Interactions, Subcellular Distribution, and Activity of p21ras GTPase Activating Protein. **Mol. Cell. Biol**. 11, 1804-1812

**16.** Rubinfeld, B., Wong, G., Bekesi, E. Wood, A. McCormick, F. and **Polakis, P. G.** 1991 A Synthetic Peptide Corresponding to a Sequence in the GTPase Activating Protein Inhibits $p21^{ras}$ Stimulation and Promotes Guanine Nucleotide Exchange. **Internatl. J. Peptide and Prot. Res**. 38, 47-53.

**17.** Rubinfeld, B., Munemitsu, S., Clark, R., Conroy, L., Watt, K., Crosier, W., McCormick, F., and **Polakis, P**. 1991 Molecular Cloning of a GTPase Activating Protein Specific for the Krev-1 Protein $p21^{rap1}$. **Cell** 65, 1033-1042.

**18.** Zhang, K. Papageorge, A., G., Martin, P., Vass, W. C., Olah, Z., **Polakis, P.**, McCormick, F. and Lowy, D, R. 1991 Heterogenous Amino Acids in RAS and Rap1A Specifying Sensitivity to GAP Proteins. **Science** 254, 1630-1634.

**19.** Martin, G., Yatani, A., Clark, R., **Polakis, P.**, Brown, A. M. and McCormick, F. 1992 GAP Domains Responsible for $p21^{ras}$-dependent Inhibition of Muscarinic Atrial $K^+$ Channel Currents. **Science** 255, 192-194.

**20.** McCormick, F., Martin, G. A., Clark, R., Bollag, G. and **Polakis, P** . 1992 Regulation of p21ras by GTPase Activating Proteins. Cold Spring Harbor **Symposia on Quantitative Biology**. Vol. 56, 237-241.

**21.** Pronk, G. B., **Polakis, P.**, Wong, G., deVries-Smits, A. M., Bos J. L. and McCormick, F. 1992 $p60^{v\text{-}src}$ Can Associate with and Phosphorylate the $p21^{ras}$ GTPase Activating Protein. **Oncogene** 7,389-394.

**22. Polakis P**. and McCormick, F. 1992 Interactions Between $p21^{ras}$ Proteins and Their GTPase Activating Proteins. In **Cancer Surveys** ( Franks, L. M., ed.) 12, 25-42.

23. Wong, G., Muller, O., Clark, R., Conroy, L., Moran, M., **Polakis,P.** and McCormick, F. 1992 Molecular coloning and nucleic acid binding properties of the GAP-associated tyrosine phosphoprotein p62. **Cell** 69, 551-558.

24. **Polakis, P.**, Rubinfeld, B. and McCormick, F. 1992 Phosphorylation of rap1GAP in vivo and by cAMP-dependent Kinase and the Cell Cycle $p34^{cdc2}$ Kinase in vitro. **J. Biol. Chem**. 267, 10780-10785.

25. McCabe, P.C., Haubrauck, H., **Polakis, P.**, McCormick, F., and Innis, M. A. 1992 Functional Interactions Between $p21^{rap1A}$ and Components of the Budding pathway of *Saccharomyces cerevisiae*. **Mol. Cell. Biol**. 12, 4084-4092.

26. Rubinfeld, B., Crosier, W.J., Albert, I., Conroy,L., Clark, R., McCormick, F. and **Polakis, P**. 1992 Localization of the rap1GAP Catalytic Domain and Sites of Phosphorylation by Mutational Analysis. **Mol. Cell . Biol**. 12, 4634-4642.

27. Ando, S., Kaibuchi, K., Sasaki, K., Hiraoka, T., Nishiyama, T., Mizuno, T., Asada, M., Nunoi, H., Matsuda, I., Matsuura, Y., **Polakis, P.**, McCormick, F. and Takai, Y. 1992 Post-translational processing of rac p21s is important both for their interaction with the GDP/GTP exchange proteins and for their activation of NADPH oxidase. **J. Biol. Chem**. 267, 25709-25713.

28. Janoueix-Lerosey, I., **Polakis, P.**, Tavitian, A. and deGunzberg, J. 1992 Regulation of the GTPase activity of the ras-related rap2 protein. **Biochem. Biophys. Res. Commun**. 189, 455-464.

29. **Polakis, P**. 1993 GAPs Specific for the rap1/Krev-1 Protein. in GTP-binding Proteins: the *ras*-superfamily. ( J.C. LaCale and F. McCormick, eds.) 445-452.

30. **Polakis, P.** and McCormick, F. 1993 Structural requirements for the interaction of $p21^{ras}$ with GAP, exchange factors, and its bological effector target. **J. Biol Chem**. 268, 9157-9160.

31. Rubinfeld, B., Souza, B. Albert, I., Muller, O., Chamberlain, S., Masiarz, F., Munemitsu, S. and **Polakis, P.** 1993 Association of the APC gene product with beta- catenin. **Science** 262, 1731-1734.

32. Weiss, J., Rubinfeld, B., **Polakis, P.**, McCormick, F. Cavenee, W. A. and Arden, K. 1993 The gene for human rap1-GTPase activating protein (rap1GAP) maps to chromosome 1p35-1p36.1. **Cytogenet. Cell Genet**. 66, 18-21.

33. Sato, K. Y., **Polakis, P.**, Haubruck, H., Fasching, C. L., McCormick, F. and Stanbridge, E. J. 1994 Analysis of the tumor suppressor acitvity of the K-rev gene in human tumor cell lines. **Cancer Res**. 54, 552-559.

34. Janoueix-Lerosey, I., Fontenay, M., Tobelem, G., Tavitian, A., **Polakis, P.** and DeGunzburg, J. 1994 Phosphorylation of rap1GAP during the cell cycle. **Biochem. Biophys. Res. Commun**. 202, 967-975

35. Munemitsu, S., Souza, B., Mueller, O., Albert, I., Rubinfeld, B., and **Polakis, P.** 1994 The APC gene product associates with microtubules in vivo and affects their assembly in vitro. **Cancer Res**. 54, 3676-3681.

36. Rubinfeld, B. and **Polakis, P.** 1995 Purification of baculovirus produced rap1GAP. **Methods Enz.** 255,31

37. **Polakis, P.** 1995 Mutations in the APC gene and their implications for protein structure and function. **Current Opinions in Genetics and Development** 5, 66-71

38. Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S. and **Polakis P.** 1995 The APC protein and E-cadherin form similar but independent complexes with α-catenin, β-catenin and Plakoglobin. **J. Biol. Chem.** 270, 5549-5555

39. Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B., and **Polakis, P.** 1995 Regulation of intracellular β-catenin levels by the APC tumor suppressor gene. **Proc. Natl. Acad. Sci.** 92, 3046-3050.

40. Lock, P., Fumagalli, S., **Polakis, P.** McCormick, F. and Courtneidge, S. A. 1996 The human p62 cDNA encodes Sam68 and not the rasGAP-associated p62 protein. **Cell** 84, 23-24.

41. Papkoff, J., Rubinfeld, B., Schryver, B. and **Polakis, P.** 1996 Wnt-1 regulates free pools of catenins and stabilizes APC-catenin complexes. **Mol. Cell. Biol.** 16, 2128-2134.

42. Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S. and **Polakis, P.** 1996 Binding of GSK3β to the APC-β-catenin complex and regulation of complex assembly. **Science** 272, 1023-1026.

43. Munemitsu, S., Albert, I., Rubinfeld, B. and **Polakis, P.** 1996 Deletion of amino-terminal structure stabilizes β-catenin in vivo and promotes the hyperphosphorylation of the APC tumor suppressor protein. **Mol. Cell. Biol.**16, 4088-4094.

44. Hart, M. J., Callow, M. G., Sousa, B. and **Polakis P.** 1996 IQGAP1, a calmodulin binding protein witha rasGAP related domain, is a potential effector for cdc42Hs. **EMBO J.** 15, 2997-3005.

45. Nathke, I. S., Adams, C. L., **Polakis, P.**, Sellin, J. and Nelson, W. J. 1996 The adenomatous polyposis coli (APC) tumor suppressor protein is localized to plasma membrane sites involved in active epithelial cell migration. **J. Cell. Biol.** 134, 165-180.

46. Hart, M. J., Sharma, S., elMasry, N., Qui, R-G., McCabe, P., **Polakis, P.** and Bollag, G. 1996 Identification of a novel guanine nucleotide exchange factor for the rho GTPase. **J. Biol. Chem.** 271, 25452.

47. Thomas JE, Smith M, Rubinfeld B, Gutowski M, Beckmann RP, and **Polakis P.** 1996 Subcellular localization and analysis of apparent 180-kDa and 220-kDa proteins of the breast cancer susceptibility gene, BRCA1. **J. Biol. Chem.** 1996 271, 28630-28635

48. Hayashi, S., Rubinfeld, B., Souza, B., **Polakis, P.**, Wieschaus, E., and Levine, A. 1997 A Drosophila homolog of the tumor suppressor adenomatous polyposis coli

down-regulates β -catenin but its zygotic expression is not essential for the regulation of armadillo. **Proc. Natl. Acad. Sci.** 94, 242-247.

49. Vleminckx, K., Rubinfeld, B., **Polakis, P.** and Gumbiner, B. 1997 The APC tumor suppressor protein induces a new axis in Xenopus embryos. **J. Cell. Biol.** 136, 411-420.

50. Rubinfeld, B., Robbins, P., El-Gamil, M., Albert, I., Porfiri, P. and **Polakis, P.** 1997 Stabilization of β-catenin by genetic defects in melanoma cell lines. **Science** 275, 1790-1792.

51. **Polakis, P.** The adenomatous polyposis coli (APC) tumor suppressor. 1997 **Biochem. Biophys. Acta**, 1332, F127-F147.

52. Rubinfeld, B., Albert, I., Porfiri, E., Munemitsu, S., and **Polakis, P** 1997 Loss of β-catenin regulation by the APC tumor suppressor protein correlates with loss of structure due to common somatic mutations of the gene. **Cancer Res.** 57, 4624-4630.

53. Porfiri, E., Rubinfeld, B., Albert, I., Hovanes. K., Waterman, M., and **Polakis, P.** 1997 Induction of a β-catenin-LEF-1 complex by wnt-1 and transforming mutants of β-catenin. **Oncogene** 15, 2833-2839.

54. Thomas JE, Smith M, Tonkinson JL, Rubinfeld B, and **Polakis P.**, 1997 Induction of phosphorylation on BRCA1 during the cell cycle and after DNA damage. **Cell Growth Differ.** 8, 801-809.

55. Hart, M., de los Santos, R., Albert, I., Rubinfeld, B., and **Polakis P.**, 1998 Down regulation of β-catenin by human Axin and its association with the adenomatous polyposis coli (APC) tumor suppressor, β-catenin and glycogen synthase kinase 3β. **Current Biology 8,** 573-581.

56. **Polakis, P.** 1998 The oncogenic activation of β-catenin. **Current Opinions in Genetics and Development** 9, 15-21

57. Matt Hart, Jean-Paul Concordet, Irina Lassot, Iris Albert, Rico del los Santos, Herve Durand, Christine Perret, Bonnee Rubinfled, Florence Margottin, Richard Benarous and **Paul Polakis**. 1999 The F-box protein β–TrCP associates with phosphorylated β–catenin and regulates its activity in the cell. **Current Biology** 9, 207-10.

58. Howard C. Crawford, Barbara M. Fingleton, Bonnee Rubinfeld, **Paul Polakis** and Lynn M. Matrisian 1999 The metalloproteinase matrilysin is a target of β–catenin transactivation in intestinal tumours. **Oncogene** 18, 2883-91.

59. Meng J, Glick JL, **Polakis P**, Casey PJ. 1999 Functional interaction between Galpha(z) and Rap1GAP suggests a novel form of cellular cross-talk. **J Biol Chem.** 17, 36663-9

60. Vijayasurian Easwaran, Virginia Song, **Paul Polakis** and Steve Byers 1999 The ubiquitin-proteosome pathway and serine kinase activity modulate APC mediated regulation of β-catenin-LEF signaling. **J. Biol. Chem.** 274(23):16641-5.

**61 Polakis P**, Hart M and Rubinfeld B. 1999 Defects in the regulation of beta-catenin in colorectal cancer. **Adv Exp Med Biol.** 470, 23-32

62 Shen Z, Batzer A, Koehler JA, **Polakis P**, Schlessinger J, Lydon NB, Moran MF. 1999 Evidence for SH3 domain directed binding and phosphorylation of Sam68 by Src. **Oncogene.** 18, 4647-53

64. Thomas GM, Frame S, Goedert M, Nathke I, **Polakis P**, Cohen P. 1999 A GSK3- binding peptide from FRAT1 selectively inhibits the GSK3-catalysed phosphorylation of axin and beta-catenin. **FEBS Lett.** 458, 247-51.

65. Peifer M, **Polakis P.** 2000 Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus. **Science** 287,1606-9.

66. **Polakis P.** 2000 Wnt signaling and cancer. **Genes Dev**;14, 1837-1851.

67. Spink KE, **Polakis P**, Weis WI 2000 Structural basis of the Axin-adenomatous polyposis coli interaction. **EMBO J** 19, 2270-2279.

68. Szeto , W., Jiang, W., Tice, D.A., Rubinfeld, B., Hollingshead, P.G., Fong, S.E., Dugger, D.L., Pham, T., Yansura, D.E., Wong, T.A., Grimaldi, J.C., Corpuz, R.T., Singh J.S., Frantz, G.D., Devaux, B., Crowley, C.W., Schwall, R.H., Eberhard, D.A., Rastelli, L., **Polakis, P.** and Pennica, D. 2001 Overexpression of the Retinoic Acid-Responsive Gene Stra6 in Human Cancers and its Synergistic Induction by Wnt-1 and Retinoic Acid. **Cancer Res** 61, 4197-4204.

69. Rubinfeld B, Tice DA, **Polakis P.** 2001 Axin dependent phosphorylation of the adenomatous polyposis coli protein mediated by casein kinase 1 epsilon. **J Biol Chem** 276, 39037-39045.

70. **Polakis P.** 2001 More than one way to skin a catenin. **Cell** 2001 105, 563-566.

71. Tice DA, Soloviev I, **Polakis P.** 2002 Activation of the Wnt Pathway Interferes withSerum Response Element-driven Transcription of Immediate Early Genes. **J Biol. Chem.** 277, 6118-6123.

72. Tice DA, Szeto W, Soloviev I, Rubinfeld B, Fong SE, Dugger DL, Winer J,

Williams PM, Wieand D, Smith V, Schwall RH, Pennnica D, **Polakis P.** 2002 Synergistic activation of tumor antigens by wnt-1 signaling and retinoic acid revealed by gene expression profiling. **J Biol Chem.** 277,14329-14335.

73. **Polakis, P.** 2002 Casein kinase I: A wnt'er of disconnect. **Curr. Biol.** 12, R499.

74. Mao,W. , Luis, E., Ross, S., Silva, J., Tan, C., Crowley, C., Chui, C., Franz, G., Senter, P., Koeppen, H., **Polakis, P.** 2004 EphB2 as a therapeutic antibody drug target for the treatment of colorectal cancer. **Cancer Res.** 64, 781-788.

75. Shibamoto, S., Winer, J., Williams, M., Polakis, P. 2003 A Blockade in Wnt signaling is activated following the differentiation of F9 teratocarcinoma cells. **Exp. Cell Res.** 29211-20.

76. Zhang Y, Eberhard DA, Frantz GD, Dowd P, Wu TD, Zhou Y, Watanabe C, Luoh SM, **Polakis P,** Hillan KJ, Wood WI, Zhang Z. 2004 GEPIS--quantitative gene expression profiling in normal and cancer tissues. **Bioinformatics,** April 8

# IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

| | | |
|---|---|---|
| Applicant | : | Ashkenazi et al. |
| App. No. | : | 09/903,925 |
| Filed | : | July 11, 2001 |
| For | : | SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME |
| Examiner | : | Hamud, Fozia M |

Group Art Unit 1647

CERTIFICATE OF EXPRESS MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to Commissioner of Patents, Washington D.C. 20231 on:

_____
(Date)

_____

Commissioner of Patents
P.O. Box 1450
Alexandria, VA 22313-1450

## DECLARATION OF AVI ASHKENAZI, Ph.D UNDER 37 C.F.R. § 1.132

I, Avi Ashkenazi, Ph.D. declare and say as follows: -

1. I am Director and Staff Scientist at the Molecular Oncology Department of Genentech, Inc., South San Francisco, CA 94080.

2. I joined Genentech in 1988 as a postdoctoral fellow. Since then, I have investigated a variety of cellular signal transduction mechanisms, including apoptosis, and have developed technologies to modulate such mechanisms as a means of therapeutic intervention in cancer and autoimmune disease. I am currently involved in the investigation of a series of secreted proteins over-expressed in tumors, with the aim to identify useful targets for the development of therapeutic antibodies for cancer treatment.

3. My scientific Curriculum Vitae, including my list of publications, is attached to and forms part of this Declaration (Exhibit A).

4. Gene amplification is a process in which chromosomes undergo changes to contain multiple copies of certain genes that normally exist as a single copy, and is an important factor in the pathophysiology of cancer. Amplification of certain genes (e.g., Myc or Her2/Neu)

gives cancer cells a growth or survival advantage relative to normal cells, and might also provide a mechanism of tumor cell resistance to chemotherapy or radiotherapy.

5. If gene amplification results in over-expression of the mRNA and the corresponding gene product, then it identifies that gene product as a promising target for cancer therapy, for example by the therapeutic antibody approach. Even in the absence of over-expression of the gene product, amplification of a cancer marker gene - as detected, for example, by the reverse transcriptase TaqMan® PCR or the fluorescence *in situ* hybridization (FISH) assays -is useful in the diagnosis or classification of cancer, or in predicting or monitoring the efficacy of cancer therapy. An increase in gene copy number can result not only from intrachromosomal changes but also from chromosomal aneuploidy. It is important to understand that detection of gene amplification can be used for cancer diagnosis even if the determination includes measurement of chromosomal aneuploidy. Indeed, as long as a significant difference relative to normal tissue is detected, it is irrelevant if the signal originates from an increase in the number of gene copies per chromosome and/or an abnormal number of chromosomes.

6. I understand that according to the Patent Office, absent data demonstrating that the increased copy number of a gene in certain types of cancer leads to increased expression of its product, gene amplification data are insufficient to provide substantial utility or well established utility for the gene product (the encoded polypeptide), or an antibody specifically binding the encoded polypeptide. However, even when amplification of a cancer marker gene does not result in significant over-expression of the corresponding gene product, this very absence of gene product over-expression still provides significant information for cancer diagnosis and treatment. Thus, if over-expression of the gene product does not parallel gene amplification in certain tumor types but does so in others, then parallel monitoring of gene amplification and gene product over-expression enables more accurate tumor classification and hence better determination of suitable therapy. In addition, absence of over-expression is crucial information for the practicing clinician. If a gene is amplified but the corresponding gene product is not over-expressed, the clinician accordingly will decide not to treat a patient with agents that target that gene product.

7. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so

made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful statements may jeopardize the validity of the application or any patent issued thereon.

By: Avi Ashkenazi
Avi Ashkenazi, Ph.D.

Date: 9/15/03

SV 455281 v1
9/12/03 3:06 PM (39780.7000)

# EXHIBIT A

# CURRICULUM VITAE

**Avi Ashkenazi**

July 2003

**Personal:**


| | |
|---|---|
| Date of birth: | 29 November, 1956 |
| Address: | 1456 Tarrytown Street, San Mateo, CA 94402 |
| Phone: | (650) 578-9199 (home); (650) 225-1853 (office) |
| Fax: | (650) 225-6443 (office) |
| Email: | aa@gene.com |


**Education:**

| | |
|---|---|
| 1983: | B.S. in Biochemistry, with honors, Hebrew University, Israel |
| 1986: | Ph.D. in Biochemistry, Hebrew University, Israel |

**Employment:**

| | |
|---|---|
| 1983-1986: | Teaching assistant, undergraduate level course in Biochemistry |
| 1985-1986: | Teaching assistant, graduate level course on Signal Transduction |
| 1986 - 1988: | Postdoctoral fellow, Hormone Research Dept., UCSF, and Developmental Biology Dept., Genentech, Inc., with J. Ramachandran |
| 1988 - 1989: | Postdoctoral fellow, Molecular Biology Dept., Genentech, Inc., with D. Capon |
| 1989 - 1993: | Scientist, Molecular Biology Dept., Genentech, Inc. |
| 1994 -1996: | Senior Scientist, Molecular Oncology Dept., Genentech, Inc. |
| 1996-1997: | Senior Scientist and Interim director, Molecular Oncology Dept., Genentech, Inc. |
| 1997-1990: | Senior Scientist and preclinical project team leader, Genentech, Inc. |
| 1999 -2002: | Staff Scientist in Molecular Oncology, Genentech, Inc. |
| 2002-present: | Staff Scientist and Director in Molecular Oncology, Genentech, Inc. |

**Awards:**

| | |
|---|---|
| 1988: | First prize, The Boehringer Ingelheim Award |

**Editorial:**

Editorial Board Member: Current Biology
Associate Editor, Clinical Cancer Research.
Associate Editor, Cancer Biology and Therapy.

**Refereed papers:**

1. Gertler, A., Ashkenazi, A., and Madar, Z. Binding sites for human growth hormone and ovine and bovine prolactins in the mammary gland and liver of the lactating cow. *Mol. Cell. Endocrinol.* **34**, 51-57 (1984).
2. Gertler, A., Shamay, A., Cohen, N., Ashkenazi, A., Friesen, H., Levanon, A., Gorecki, M., Aviv, H., Hadari, D., and Vogel, T. Inhibition of lactogenic activities of ovine prolactin and human growth hormone (hGH) by a novel form of a modified recombinant hGH. *Endocrinology* **118**, 720-726 (1986).
3. Ashkenazi, A., Madar, Z., and Gertler, A. Partial purification and characterization of bovine mammary gland prolactin receptor. *Mol. Cell. Endocrinol.* **50**, 79-87 (1987).
4. Ashkenazi, A., Pines, M., and Gertler, A. Down-regulation of lactogenic hormone receptors in Nb2 lymphoma cells by cholera toxin. *Biochemistry Internatl.* **14**, 1065-1072 (1987).
5. Ashkenazi, A., Cohen, R., and Gertler, A. Characterization of lactogen receptors in lactogenic hormone-dependent and independent Nb2 lymphoma cell lines. *FEBS Lett.* **210**, 51-55 (1987).
6. Ashkenazi, A., Vogel, T., Barash, I., Hadari, D., Levanon, A., Gorecki, M., and Gertler, A. Comparative study on in vitro and in vivo modulation of lactogenic and somatotropic receptors by native human growth hormone and its modified recombinant analog. *Endocrinology* **121**, 414-419 (1987).
7. Peralta, E., Winslow, J., Peterson, G., Smith, D., Ashkenazi, A., Ramachandran, J., Schimerlik, M., and Capon, D. Primary structure and biochemical properties of an M2 muscarinic receptor. *Science* **236**, 600-605 (1987).
8. Peralta, E. Ashkenazi, A., Winslow, J., Smith, D., Ramachandran, J., and Capon, D. J. Distinct primary structures, ligand-binding properties and tissue-specific expression of four human muscarinic acetylcholine receptors. *EMBO J.* **6**, 3923-3929 (1987).
9. Ashkenazi, A., Winslow, J., Peralta, E., Peterson, G., Schimerlik, M., Capon, D., and Ramachandran, J. An M2 muscarinic receptor subtype coupled to both adenylyl cyclase and phosphoinositide turnover. *Science* **238**, 672-675 (1987).

10. Pines, M., Ashkenazi, A., Cohen-Chapnik, N., Binder, L., and Gertler, A. Inhibition of the proliferation of Nb2 lymphoma cells by femtomolar concentrations of cholera toxin and partial reversal of the effect by 12-o-tetradecanoyl-phorbol-13-acetate. *J. Cell. Biochem.* **37**, 119-129 (1988).

11. Peralta, E. Ashkenazi, A., Winslow, J. Ramachandran, J., and Capon, D. Differential regulation of PI hydrolysis and adenylyl cyclase by muscarinic receptor subtypes. *Nature* **334**, 434-437 (1988).

12. Ashkenazi., A. Peralta, E., Winslow, J., Ramachandran, J., and Capon, D. Functionally distinct G proteins couple different receptors to PI hydrolysis in the same cell. *Cell* **56**, 487-493 (1989).

13. Ashkenazi, A., Ramachandran, J., and Capon, D. Acetylcholine analogue stimulates DNA synthesis in brain-derived cells via specific muscarinic acetylcholine receptor subtypes. *Nature* **340**, 146-150 (1989).

14. Lammare, D., Ashkenazi, A., Fleury, S., Smith, D., Sekaly, R., and Capon, D. The MHC-binding and gp120-binding domains of CD4 are distinct and separable. *Science* **245**, 743-745 (1989).

15. Ashkenazi., A., Presta, L., Marsters, S., Camerato, T., Rosenthal, K., Fendly, B., and Capon, D. Mapping the CD4 binding site for human immunodefficiency virus type 1 by alanine-scanning mutagenesis. *Proc. Natl. Acad. Sci. USA.* **87**, 7150-7154 (1990).

16. Chamow, S., Peers, D., Byrn, R., Mulkerrin, M., Harris, R., Wang, W., Bjorkman, P., Capon, D., and Ashkenazi, A. Enzymatic cleavage of a CD4 immunoadhesin generates crystallizable, biologically active Fd-like fragments. *Biochemistry* **29**, 9885-9891 (1990).

17. Ashkenazi, A., Smith, D., Marsters, S., Riddle, L., Gregory, T., Ho, D., and Capon, D. Resistance of primary isolates of human immunodefficiency virus type 1 to soluble CD4 is independent of CD4-rgp120 binding affinity. *Proc. Natl. Acad. Sci. USA.* **88**, 7056-7060 (1991).

18. Ashkenazi, A., Marsters, S., Capon, D., Chamow, S., Figari., I., Pennica, D., Goeddel., D., Palladino, M., and Smith, D. Protection against endotoxic shock by a tumor necrosis factor receptor immunoadhesin. *Proc. Natl. Acad. Sci. USA.* **88**, 10535-10539 (1991).

19. Moore, J., McKeating, J., Huang, Y., Ashkenazi, A., and Ho, D. Virions of primary HIV-1 isolates resistant to sCD4 neutralization differ in sCD4 affinity and glycoprotein gp120 retention from sCD4-sensitive isolates. *J. Virol.* **66**, 235-243 (1992).

20. Jin, H., Oksenberg, D., Ashkenazi, A., Peroutka, S., Duncan, A., Rozmahel., R., Yang, Y., Mengod, G., Palacios, J., and O'Dowd, B. Characterization of the human 5-hydroxytryptamine$_{1B}$ receptor. *J. Biol. Chem.* **267**, 5735-5738 (1992).

21. Marsters, A., Frutkin, A., Simpson, N., Fendly, B. and Ashkenazi, A. Identification of cysteine-rich domains of the type 1 tumor necrosis receptor involved in ligand binding. *J. Biol. Chem.* **267**, 5747-5750 (1992).

22. Chamow, S., Kogan, T., Peers, D., Hastings, R., Byrn, R., and Ashkenazi, A. Conjugation of sCD4 without loss of biological activity via a novel carbohydrate-directed cross-linking reagent. *J. Biol. Chem.* **267**, 15916-15922 (1992).

23. Oksenberg, D., Marsters, A., O'Dowd, B., Jin, H., Havlik, S., Peroutka, S., and Ashkenazi, A. A single amino-acid difference confers major pharmacologic variation between human and rodent 5-HT$_{1B}$ receptors. *Nature* **360**, 161-163 (1992).

24. Haak-Frendscho, M., Marsters, S., Chamow, S., Peers, D., Simpson, N., and Ashkenazi, A. Inhibition of interferon γ by an interferon γ receptor immunoadhesin. *Immunology* **79**, 594-599 (1993).

25. Penica, D., Lam, V., Weber, R., Kohr, W., Basa, L., Spellman, M., Ashkenazi, Shire, S., and Goeddel, D. Biochemical characterization of the extracellular domain of the 75-kd tumor necrosis factor receptor. *Biochemistry* **32**, 3131-3138. (1993).

26. Barfod, L., Zheng, Y., Kuang, W., Hart, M., Evans, T., Cerione, R., and Ashkenazi, A. Cloning and expression of a human CDC42 GTPase Activating Protein reveals a functional SH3-binding domain. *J. Biol. Chem.* **268**, 26059-26062 (1993).

27. Chamow, S., Zhang, D., Tan, X., Mhtre, S., Marsters, S., Peers, D., Byrn, R., Ashkenazi, A., and Yunghans, R. A humanized bispecific immunoadhesin-antibody that retargets CD3+ effectors to kill HIV-1-infected cells. *J. Immunol.* **153**, 4268-4280 (1994).

28. Means, R., Krantz, S., Luna, J., Marsters, S., and Ashkenazi, A. Inhibition of murine erythroid colony formation in vitro by iterferon γ and correction by interferon γ receptor immunoadhesin. *Blood* **83**, 911-915 (1994).

29. Haak-Frendscho, M., Marsters, S., Mordenti, J., Gillet, N., Chen, S., andAshkenazi, A. Inhibition of TNF by a TNF receptor immunoadhesin: comparison with an anti-TNF mAb. *J. Immunol.* **152**, 1347-1353 (1994).

30. Chamow, S., Kogan, T., Venuti, M., Gadek, T., Peers, D., Mordenti, J., Shak, S., and Ashkenazi, A. Modification of CD4 immunoadhesin with monomethoxy-PEG aldehyde via reductive alkilation. *Bioconj. Chem.* **5**, 133-140 (1994).

31. Jin, H., Yang, R., Marsters, S., Bunting, S., Wurm, F., Chamow, S., and Ashkenazi, A. Protection against rat endotoxic shock by p55 tumor necrosis factor (TNF) receptor immunoadhesin: comparison to anti-TNF monoclonal antibody. *J. Infect. Diseases* **170**, 1323-1326 (1994).

32. Beck, J., Marsters, S., Harris, R., Ashkenazi, A., and Chamow, S. Generation of soluble interleukin-1 receptor from an immunoadhesin by specific cleavage. *Mol. Immunol.* **31**, 1335-1344 (1994).

33. Pitti, B., Marsters, M., Haak-Frendscho, M., Osaka, G., Mordenti, J., Chamow, S., and Ashkenazi, A. Molecular and biological properties of an interleukin-1 receptor immunoadhesin. *Mol. Immunol.* **31**, 1345-1351 (1994).

34. Oksenberg, D., Havlik, S., Peroutka, S., and Ashkenazi, A. The third intracellular loop of the 5-HT2 receptor specifies effector coupling. *J. Neurochem.* **64**, 1440-1447 (1995).

35. Bach, E., Szabo, S., Dighe, A., Ashkenazi, A., Aguet, M., Murphy, K., and Schreiber, R. Ligand-induced autoregulation of IFN-γ receptor β chain expression in T helper cell subsets. *Science* **270**, 1215-1218 (1995).

36. Jin, H., Yang, R., Marsters, S., Ashkenazi, A., Bunting, S., Marra, M., Scott, R., and Baker, J. Protection against endotoxic shock by bactericidal/permeability-increasing protein in rats. *J. Clin. Invest.* **95**, 1947-1952 (1995).

37. Marsters, S., Penica, D., Bach, E., Schreiber, R., and Ashkenazi, A. Interferon γ signals via a high-affinity multisubunit receptor complex that contains two types of polypeptide chain. *Proc. Natl. Acad. Sci. USA.* **92**, 5401-5405 (1995).

38. Van Zee, K., Moldawer, L., Oldenburg, H., Thompson, W., Stackpole, S., Montegut, W., Rogy, M., Meschter, C., Gallati, H., Schiller, C., Richter, W., Loetcher, H., Ashkenazi, A., Chamow, S., Wurm, F., Calvano, S., Lowry, S., and Lesslauer, W. Protection against lethal E. coli bacteremia in baboons by pretreatment with a 55-kDa TNF receptor-Ig fusion protein, Ro45-2081. *J. Immunol.* **156**, 2221-2230 (1996).

39. Pitti, R., Marsters, S., Ruppert, S., Donahue, C., Moore, A., and Ashkenazi, A. Induction of apoptosis by Apo-2 Ligand, a new member of the tumor necrosis factor cytokine family. *J. Biol. Chem.* **271**, 12687-12690 (1996).

40. Marsters, S., Pitti, R., Donahue, C., Rupert, S., Bauer, K., and Ashkenazi, A. Activation of apoptosis by Apo-2 ligand is independent of FADD but blocked by CrmA. *Curr. Biol.* **6**, 1669-1676 (1996).

41. Marsters, S., Skubatch, M., Gray, C., and Ashkenazi, A. Herpesvirus entry mediator, a novel member of the tumor necrosis factor receptor family, activates the NF-κB and AP-1 transcription factors. *J. Biol. Chem.* **272**, 14029-14032 (1997).

42. Sheridan, J., Marsters, S., Pitti, R., Gurney, A., Skubatch, M., Baldwin, D., Ramakrishnan, L., Gray, C., Baker, K., Wood, W.I., Goddard, A., Godowski, P., and Ashkenazi, A. Control of TRAIL-induced apoptosis by a family of signaling and decoy receptors. *Science* **277**, 818-821 (1997).

43. Marsters, S., Sheridan, J., Pitti, R., Gurney, A., Skubatch, M., Balswin, D., Huang, A., Yuan, J., Goddard, A., Godowski, P., and Ashkenazi, A. A novel receptor for Apo2L/TRAIL contains a truncated death domain. *Curr. Biol.* **7**, 1003-1006 (1997).

44. Marsters, A., Sheridan, J., Pitti, R., Brush, J., Goddard, A., and Ashkenazi, A. Identification of a ligand for the death-domain-containing receptor Apo3. *Curr. Biol.* **8**, 525-528 (1998).

45. Rieger, J., Naumann, U., Glaser, T., Ashkenazi, A., and Weller, M. Apo2 ligand: a novel weapon against malignant glioma? *FEBS Lett.* **427**, 124-128 (1998).

46. Pender, S., Fell, J., Chamow, S., Ashkenazi, A., and MacDonald, T. A p55 TNF receptor immunoadhesin prevents T cell mediated intestinal injury by inhibiting matrix metalloproteinase production. *J. Immunol.* **160**, 4098-4103 (1998).

47. Pitti, R., Marsters, S., Lawrence, D., Roy, Kischkel, F., M., Dowd, P., Huang, A., Donahue, C., Sherwood, S., Baldwin, D., Godowski, P., Wood, W., Gurney, A., Hillan, K., Cohen, R., Goddard, A., Botstein, D., and Ashkenazi, A. Genomic amplification of a decoy receptor for Fas ligand in lung and colon cancer. *Nature* **396**, 699-703 (1998).

48. Mori, S., Marakami-Mori, K., Nakamura, S., Ashkenazi, A., and Bonavida, B. Sensitization of AIDS Kaposi's sarcoma cells to Apo-2 ligand-induced apoptosis by actinomycin D. *J. Immunol.* **162**, 5616-5623 (1999).

49. Gurney, A. Marsters, S., Huang, A., Pitti, R., Mark, M., Baldwin, D., Gray, A., Dowd, P., Brush, J., Heldens, S., Schow, P., Goddard, A., Wood, W., Baker, K., Godowski, P., and Ashkenazi, A. Identification of a new member of the tumor necrosis factor family and its receptor, a human ortholog of mouse GITR. *Curr. Biol.* **9**, 215-218 (1999).

50. Ashkenazi, A., Pai, R., Fong, s., Leung, S., Lawrence, D., Marsters, S., Blackie, C., Chang, L., McMurtrey, A., Hebert, A., DeForge, L., Khoumenis, I., Lewis, D., Harris, L., Bussiere, J., Koeppen, H., Shahrokh, Z., and Schwall, R. Safety and anti-tumor activity of recombinant soluble Apo2 ligand. *J. Clin. Invest.* **104**, 155-162 (1999).

51. Chuntharapai, A., Gibbs, V., Lu, J., Ow, A., Marsters, S., Ashkenazi, A., De Vos, A., Kim, K.J. Determination of residues involved in ligand binding and signal transmissiion in the human IFN-α receptor 2. *J. Immunol.* **163**, 766-773 (1999).

52. Johnsen, A.-C., Haux, J., Steinkjer, B., Nonstad, U., Egeberg, K., Sundan, A., Ashkenazi, A., and Espevik, T. Regulation of Apo2L/TRAIL expression in NK cells – involvement in NK cell-mediated cytotoxicity. *Cytokine* **11**, 664-672 (1999).

53. Roth, W., Isenmann, S., Naumann, U., Kugler, S., Bahr, M., Dichgans, J., Ashkenazi, A., and Weller, M. Eradication of intracranial human malignant glioma xenografts by Apo2L/TRAIL. *Biochem. Biophys. Res. Commun.* **265**, 479-483 (1999).

54. Hymowitz, S.G., Christinger, H.W., Fuh, G., Ultsch, M., O'Connell, M., Kelley, R.F., Ashkenazi, A. and de Vos, A.M. Triggering Cell Death: The Crystal Structure of Apo2L/TRAIL in a Complex with Death Receptor 5. *Molec. Cell* **4**, 563–571 (1999).

55. Hymowitz, S.G., O'Connel, M.P., Utsch, M.H., Hurst, A., Totpal, K., Ashkenazi, A., de Vos, A.M., Kelley, R.F. A unique zinc-binding site revealed by a high-resolution X-ray structure of homotrimeric Apo2L/TRAIL. *Biochemistry* **39**, 633-640 (2000).

56. Zhou, Q., Fukushima, P., DeGraff, W., Mitchell, J.B., Stetler-Stevenson, M., Ashkenazi, A., and Steeg, P.S. Radiation and the Apo2L/TRAIL apoptotic pathway preferentially inhibit the colonization of premalignant human breast cancer cells overexpressing cyclin D1. *Cancer Res.* **60**, 2611-2615 (2000).

57. Kischkel, F.C., Lawrence, D. A., Chuntharapai, A., Schow, P., Kim, J., and Ashkenazi, A. Apo2L/TRAIL-dependent recruitment of endogenous FADD and Caspase-8 to death receptors 4 and 5. *Immunity* **12**, 611-620 (2000).

58. Yan, M., Marsters, S.A., Grewal, I.S., Wang, H., *Ashkenazi, A., and *Dixit, V.M. Identification of a receptor for BlyS demonstrates a crucial role in humoral immunity. *Nature Immunol.* **1**, 37-41 (2000).

59. Marsters, S.A., Yan, M., Pitti, R.M., Haas, P.E., Dixit, V.M., and Ashkenazi, A. Interaction of the TNF homologues BLyS and APRIL with the TNF receptor homologues BCMA and TACI. *Curr. Biol.* **10**, 785-788 (2000).

60. Kischkel, F.C., and Ashkenazi, A. Combining enhanced metabolic labeling with immunoblotting to detect interactions of endogenous cellular proteins. *Biotechniques* **29**, 506-512 (2000).

61. Lawrence, D., Shahrokh, Z., Marsters, S., Achilles, K., Shih, D. Mounho, B., Hillan, K., Totpal, K. DeForge, L., Schow, P., Hooley, J., Sherwood, S., Pai, R., Leung, S., Khan, L., Gliniak, B., Bussiere, J., Smith, C., Strom, S., Kelley, S., Fox, J., Thomas, D., and Ashkenazi, A. Differential hepatocyte toxicity of recombinant Apo2L/TRAIL versions. *Nature Med.* **7**, 383-385 (2001).

62. Chuntharapai, A., Dodge, K., Grimmer, K., Schroeder, K., Martsters, S.A., Koeppen, H., Ashkenazi, A., and Kim, K.J. Isotype-dependent inhibition of tumor growth in vivo by monoclonal antibodies to death receptor 4. *J. Immunol.* **166**, 4891-4898 (2001).

63. Pollack, I.F., Erff, M., and Ashkenazi, A. Direct stimulation of apoptotic signaling by soluble Apo2L/tumor necrosis factor-related apoptosis-inducing ligand leads to selective killing of glioma cells. *Clin. Cancer Res.* **7**, 1362-1369 (2001).

64. Wang, H., Marsters, S.A., Baker, T., Chan, B., Lee, W.P., Fu, L., Tumas, D., Yan, M., Dixit, V.M., *Ashkenazi, A., and *Grewal, I.S. TACI-ligand interactions are required for T cell activation and collagen-induced arthritis in mice. *Nature Immunol.* 2, 632-637 (2001).

65. Kischkel, F.C., Lawrence, D. A., Tinel, A., Virmani, A., Schow, P., Gazdar, A., Blenis, J., Arnott, D., and Ashkenazi, A. Death receptor recruitment of endogenous caspase-10 and apoptosis initiation in the absence of caspase-8. *J. Biol. Chem.* 276, 46639-46646 (2001).

66. LeBlanc, H., Lawrence, D.A., Varfolomeev, E., Totpal, K., Morlan, J., Schow, P., Fong, S., Schwall, R., Sinicropi, D., and Ashkenazi, A Tumor cell resistance to death receptor induced apoptosis through mutational inactivation of the proapoptotitc Bcl-2 homolog Bax. *Nature Med.* 8, 274-281 (2002).

67. Miller, K., Meng, G., Liu, J., Hurst, A., Hsei, V., Wong, W-L., Ekert, R., Lawrence, D., Sherwood, S., DeForge, L., Gaudreault., Keller, G., Sliwkowski, M., Ashkenazi, A., and Presta, L. Design, Construction, and analyses of multivalent antibodies. *J. Immunol.* 170, 4854-4861 (2003).

68. Varfolomeev, E., Kischkel, F., Martin, F., Wanh, H., Lawrence, D., Olsson, C., Tom, L., Erickson, S., French, D., Schow, P., Grewal, I. and Ashkenazi, A. Immune system development in APRIL knockout mice. Submitted.

**Review articles:**

1. Ashkenazi, A., Peralta, E., Winslow, J., Ramachandran, J., and Capon, D., J. Functional role of muscarinic acetylcholine receptor subtype diversity. *Cold Spring Harbor Symposium on Quantitative Biology*. **LIII**, 263-272 (1988).
2. Ashkenazi, A., Peralta, E., Winslow, J., Ramachandran, J., and Capon, D. Functional diversity of muscarinic receptor subtypes in cellular signal transduction and growth. *Trends Pharmacol. Sci.* Dec Supplement, 12-21 (1989).
3. Chamow, S., Duliege, A., Ammann, A., Kahn, J., Allen, D., Eichberg, J., Byrn, R., Capon, D., Ward, R., and Ashkenazi, A. CD4 immunoadhesins in anti-HIV therapy: new developments. *Int. J. Cancer* Supplement 7, 69-72 (1992).
4. Ashkenazi, A., Capon, and D. Ward, R. Immunoadhesins. *Int. Rev. Immunol.* **10**, 217-225 (1993).
5. Ashkenazi, A., and Peralta, E. Muscarinic Receptors. In *Handbook of Receptors and Channels*. (S. Peroutka, ed.), CRC Press, Boca Raton, Vol. **I**, p. 1-27, (1994).
6. Krantz, S. B., Means, R. T., Jr., Lina, J., Marsters, S. A., and Ashkenazi, A. Inhibition of erythroid colony formation in vitro by gamma interferon. In *Molecular Biology of Hematopoiesis* (N. Abraham, R. Shadduck, A. Levine F. Takaku, eds.) Intercept Ltd. Paris, Vol. **3**, p. 135-147 (1994).
7. Ashkenazi, A. Cytokine neutralization as a potential therapeutic approach for SIRS and shock. *J. Biotechnology in Healthcare* **1**, 197-206 (1994).
8. Ashkenazi, A., and Chamow, S. M. Immunoadhesins: an alternative to human monoclonal antibodies. *Immunomethods: A companion to Methods in Enzimology* **8**, 104-115 (1995).
9. Chamow, S., and Ashkenazi, A. Immunoadhesins: Principles and Applications. *Trends Biotech.* **14**, 52-60 (1996).
10. Ashkenazi, A., and Chamow, S. M. Immunoadhesins as research tools and therapeutic agents. *Curr. Opin. Immunol.* **9**, 195-200 (1997).
11. Ashkenazi, A., and Dixit, V. Death receptors: signaling and modulation. *Science* **281**, 1305-1308 (1998).
12. Ashkenazi, A., and Dixit, V. Apoptosis control by death and decoy receptors. *Curr. Opin. Cell. Biol.* **11**, 255-260 (1999).

13. Ashkenazi, A. Chapters on Apo2L/TRAIL; DR4, DR5, DcR1, DcR2; and DcR3. Online Cytokine Handbook (www.apnet.com/cytokinereference/).
14. Ashkenazi, A. Targeting death and decoy receptors of the tumor necrosis factor superfamily. *Nature Rev. Cancer* 2, 420-430 (2002).
15. LeBlanc, H. and Ashkenazi, A. Apoptosis signaling by Apo2L/TRAIL. *Cell Death and Differentiation* 10, 66-75 (2003).
16. Almasan, A. and Ashkenazi, A. Apo2L/TRAIL: apoptosis signaling, biology, and potential for cancer therapy. Cytokine and Growth Factor Reviews 14, 337-348 (2003).

**Book:**

Antibody Fusion Proteins (Chamow, S., and Ashkenazi, A., eds., John Wiley and Sons Inc.) (1999).

**Talks:**

1. Resistance of primary HIV isolates to CD4 is independent of CD4-gp120 binding affinity. UCSD Symposium, HIV Disease: Pathogenesis and Therapy. Greenelefe, FL, March 1991.
2. Use of immuno-hybrids to extend the half-life of receptors. IBC conference on Biopharmaceutical Halflife Extension. New Orleans, LA, June 1992.
3. Results with TNF receptor Immunoadhesins for the Treatment of Sepsis. IBC conference on Endotoxemia and Sepsis. Philadelphia, PA, June 1992.
4. Immunoadhesins: an alternative to human antibodies. IBC conference on Antibody Engineering. San Diego, CA, December 1993.
5. Tumor necrosis factor receptor: a potential therapeutic for human septic shock. American Society for Microbiology Meeting, Atlanta, GA, May 1993.
6. Protective efficiacy of TNF receptor immunoadhesin vs anti-TNF monoclonal antibody in a rat model for endotoxic shock. 5th International Congress on TNF. Asilomar, CA, May 1994.
7. Interferon-γ signals via a multisubunit receptor complex that contains two types of polypeptide chain. American Association of Immunologists Conference. San Franciso, CA, July 1995.
8. Immunoadhesins: Principles and Applications. Gordon Research Conference on Drug Delivery in Biology and Medicine. Ventura, CA, February 1996.

9. Apo-2 Ligand, a new member of the TNF family that induces apoptosis in tumor cells. Cambridge Symposium on TNF and Related Cytokines in Treatment of Cancer. Hilton-Head, NC, March 1996.
10. Induction of apoptosis by Apo2 Ligand. American Society for Biochemistry and Molecular Biology, Symposium on Growth Factors and Cytokine Receptors. New Orleans, LA, June, 1996.
11. Apo2 ligand, an extracellular trigger of apoptosis. 2nd Clontech Symposium, Palo Alto, CA, October 1996.
12. Regulation of apoptosis by members of the TNF ligand and receptor families. Stanford University School of Medicine, Palo Alto, CA, December 1996.
13. Apo-3: anovel receptor that regulates cell death and inflammation. 4th International Congress on Immune Consequences of Trauma, Shock, and Sepsis. Munich, Germany, March 1997.
14. New members of the TNF ligand and receptor families that regulate apoptosis, inflammation, and immunity. UCLA School of Medicine, LA, CA, March 1997.
15. Immunoadhesins: an alternative to monoclonal antibodies. 5th World Conference on Bispecific Antibodies. Volendam, Holland, June 1997.
16. Control of Apo2L signaling. Cold Spring Harbor Laboratory Symposium on Programmed Cell Death. Cold Spring Harbor, New York. September, 1997.
17. Chairman and speaker, Apoptosis Signaling session. IBC's 4th Annual Conference on Apoptosis. San Diego, CA., October 1997.
18. Control of Apo2L signaling by death and decoy receptors. American Association for the Advancement of Science. Philladelphia, PA, February 1998.
19. Apo2 ligand and its receptors. American Society of Immunologists. San Francisco, CA, April 1998.
20. Death receptors and ligands. 7th International TNF Congress. Cape Cod, MA, May 1998.
21. Apo2L as a potential therapeutic for cancer. UCLA School of Medicine. LA, CA, June 1998.
22. Apo2L as a potential therapeutic for cancer. Gordon Research Conference on Cancer Chemotherapy. New London, NH, July 1998.
23. Control of apoptosis by Apo2L. Endocrine Society Conference, Stevenson, WA, August 1998.
24. Control of apoptosis by Apo2L. International Cytokine Society Conference, Jerusalem, Israel, October 1998.

25. Apoptosis control by death and decoy receptors. American Association for Cancer Research Conference, Whistler, BC, Canada, March 1999.
26. Apoptosis control by death and decoy receptors. American Society for Biochemistry and Molecular Biology Conference, San Francisco, CA, May 1999.
27. Apoptosis control by death and decoy receptors. Gordon Research Conference on Apoptosis, New London, NH, June 1999.
28. Apoptosis control by death and decoy receptors. Arthritis Foundation Research Conference, Alexandria GA, Aug 1999.
29. Safety and anti-tumor activity of recombinant soluble Apo2L/TRAIL. Cold Spring Harbor Laboratory Symposium on Programmed Cell Death. . Cold Spring Harbor, NY, September 1999.
30. The Apo2L/TRAIL system: therapeutic potential. American Association for Cancer Research, Lake Tahoe, NV, Feb 2000.
31. Apoptosis and cancer therapy. Stanford University School of Medicine, Stanford, CA, Mar 2000.
32. Apoptosis and cancer therapy. University of Pennsylvania School of Medicine, Philladelphia, PA, Apr 2000.
33. Apoptosis signaling by Apo2L/TRAIL. International Congress on TNF. Trondheim, Norway, May 2000.
34. The Apo2L/TRAIL system: therapeutic potential. Cap-CURE summit meeting. Santa Monica, CA, June 2000.
35. The Apo2L/TRAIL system: therapeutic potential. MD Anderson Cancer Center. Houston, TX, June 2000.
36. Apoptosis signaling by Apo2L/TRAIL. The Protein Society, 14th Symposium. San Diego, CA, August 2000.
37. Anti-tumor activity of Apo2L/TRAIL. AAPS annual meeting. Indianapolis, IN Aug 2000.
38. Apoptosis signaling and anti-cancer potential of Apo2L/TRAIL. Cancer Research Institute, UC San Francisco, CA, September 2000.
39. Apoptosis signaling by Apo2L/TRAIL. Kenote address, TNF family Minisymposium, NIH. Bethesda, MD, September 2000.
40. Death receptors: signaling and modulation. Keystone symposium on the Molecular basis of cancer. Taos, NM, Jan 2001.
41. Preclinical studies of Apo2L/TRAIL in cancer. Symposium on Targeted therapies in the treatment of lung cancer. Aspen, CO, Jan 2001.

42. Apoptosis signaling by Apo2L/TRAIL. Wiezmann Institute of Science, Rehovot, Israel, March 2001.
43. Apo2L/TRAIL: Apoptosis signaling and potential for cancer therapy. Weizmann Institute of Science, Rehovot, Israel, March 2001.
44. Targeting death receptors in cancer with Apo2L/TRAIL. Cell Death and Disease conference, North Falmouth, MA, Jun 2001.
45. Targeting death receptors in cancer with Apo2L/TRAIL. Biotechnology Organization conference, San Diego, CA, Jun 2001.
46. Apo2L/TRAIL signaling and apoptosis resistance mechanisms. Gordon Research Conference on Apoptosis, Oxford, UK, July 2001.
47. Apo2L/TRAIL signaling and apoptosis resistance mechanisms. Cleveland Clinic Foundation, Cleveland, OH, Oct 2001.
48. Apoptosis signaling by death receptors: overview. International Society for Interferon and Cytokine Research conference, Cleveland, OH, Oct 2001.
49. Apoptosis signaling by death receptors. American Society of Nephrology Conference. San Francisco, CA, Oct 2001.
50. Targeting death receptors in cancer. Apoptosis: commercial opportunities. San Diego, CA, Apr 2002.
51. Apo2L/TRAIL signaling and apoptosis resistance mechanisms. Kimmel Cancer Research Center, Johns Hopkins University, Baltimore MD. May 2002.
52. Apoptosis control by Apo2L/TRAIL. (Keynote Address) University of Alabama Cancer Center Retreat, Birmingham, Ab. October 2002.
53. Apoptosis signaling by Apo2L/TRAIL. (Session co-chair) TNF international conference. San Diego, CA. October 2002.
54. Apoptosis signaling by Apo2L/TRAIL. Swiss Institute for Cancer Research (ISREC). Lausanne, Swizerland. Jan 2003.
55. Apoptosis induction with Apo2L/TRAIL. Conference on New Targets and Innovative Strategies in Cancer Treatment. Monte Carlo. February 2003.
56. Apoptosis signaling by Apo2L/TRAIL. Hermelin Brain Tumor Center Symposium on Apoptosis. Detroit, MI. April 2003.
57. Targeting apoptosis through death receptors. Sixth Annual Conference on Targeted Therapies in the Treatment of Breast Cancer. Kona, Hawaii. July 2003.
58. Targeting apoptosis through death receptors. Second International Conference on Targeted Cancer Therapy. Washington, DC. Aug 2003.

**Issued Patents:**

1. Ashkenazi, A., Chamow, S. and Kogan, T. Carbohydrate-directed crosslinking reagents. US patent 5,329,028 (Jul 12, 1994).
2. Ashkenazi, A., Chamow, S. and Kogan, T. Carbohydrate-directed crosslinking reagents. US patent 5,605,791 (Feb 25, 1997).
3. Ashkenazi, A., Chamow, S. and Kogan, T. Carbohydrate-directed crosslinking reagents. US patent 5,889,155 (Jul 27, 1999).
4. Ashkenazi, A., APO-2 Ligand. US patent 6,030,945 (Feb 29, 2000).
5. Ashkenazi, A., Chuntharapai, A., Kim, J., APO-2 ligand antibodies. US patent 6, 046, 048 (Apr 4, 2000).
6. Ashkenazi, A., Chamow, S. and Kogan, T. Carbohydrate-directed crosslinking reagents. US patent 6,124,435 (Sep 26, 2000).
7. Ashkenazi, A., Chuntharapai, A., Kim, J., Method for making monoclonal and cross-reactive antibodies. US patent 6,252,050 (Jun 26, 2001).
8. Ashkenazi, A. APO-2 Receptor. US patent 6,342,369 (Jan 29, 2002).
9. Ashkenazi, A. Fong, S., Goddard, A., Gurney, A., Napier, M., Tumas, D., Wood, W. A-33 polypeptides. US patent 6,410,708 (Jun 25, 2002).
10. Ashkenazi, A. APO-3 Receptor. US patent 6,462,176 B1 (Oct 8, 2002).
11. Ashkenazi, A. APO-2LI and APO-3 polypeptide antibodies. US patent 6,469,144 B1 (Oct 22, 2002).
12. Ashkenazi, A., Chamow, S. and Kogan, T. Carbohydrate-directed crosslinking reagents. US patent 6,582,928B1 (Jun 24, 2003).

###
Tue Feb 12 09:55:34 2002 [BLASTP 2.2.1 [Jul-12-2001], NCBI]
/home/ruby/va/Molbio/carpenda/tempids/p1.DNA73735 (225 aa)
/home/ruby/va/Molbio/carpenda/tempids/p1.DNA73735

| Sequences producing High-scoring Segment Pairs: | | | Score | Match | Pct | E-val |
|---|---|---|---|---|---|---|
| 1 | P_AAB87585 | Human PRO1573 - Homo sapiens. | 1172 | 225 | 100 | e-127 |
| 2 | P_AAG89285 | Human secreted protein, SEQ ID NO: 405 - | 1172 | 225 | 100 | e-127 |
| 3 | P_AAM79556 | Human protein SEQ ID NO 3202 - Homo sapie | 1172 | 225 | 100 | e-127 |
| 4 | P_AAM78572 | Human protein SEQ ID NO 1234 - Homo sapie | 1172 | 225 | 100 | e-127 |
| 5 | P_AAU29202 | Human PRO polypeptide sequence #179 - Hom | 1172 | 225 | 100 | e-127 |
| 6 | P_AAY99433 | Human PRO1573 (UNQ779) amino acid sequenc | 1172 | 225 | 100 | e-127 |
| 7 | BAA95567.114 | CLDN8 - Homo sapiens | 1172 | 225 | 100 | e-127 |
| 8 | AAH20866.114 | claudin 8 - Homo sapiens | 1172 | 225 | 100 | e-127 |
| 9 | NP_036264.1 | claudin 8 - Homo sapiens | 1172 | 225 | 100 | e-127 |
| 10 | CLD8_HUMAN | Claudin-8 /pid=CAB60615.1 - homo sapiens | 1172 | 225 | 100 | e-127 |
| 11 | P_AAG89273 | Human secreted protein, SEQ ID NO: 393 - | 1162 | 223 | 99 | e-126 |
| 12 | P_AAU00678 | Human INTERCEPT 309 polypeptide - Homo sa | 1123 | 215 | 100 | e-122 |
| 13 | P_AAB93923 | Human protein sequence SEQ ID NO:13910 - | 1111 | 213 | 99 | e-120 |

Dayhoff Protein Database (Rel 78, Mar 2004)

P_AAB87585 Human PRO1573 - Homo sapiens.
Length: 225 aa
Accession: P_AAB87585;
Species: Homo sapiens.
Keywords: Human; PRO protein; mapping; patent; GENESEQ patentdb.
Patent number: WO200116318-A2.
Publication date: 08-MAR-2001.
Filing date: 24-AUG-2000; 2000WO-US023328.
Priority: 01-SEP-1999; 99WO-US020111. 15-SEP-1999; 99WO-US021090.
07-DEC-1999; 99US-0169495P. 09-DEC-1999; 99US-0170262P.
11-JAN-2000; 2000US-0175481P. 18-FEB-2000; 2000WO-US004341.
18-FEB-2000; 2000WO-US004342. 22-FEB-2000; 2000WO-US004414.
01-MAR-2000; 2000WO-US005601. 03-MAR-2000; 2000US-0187202P.
21-MAR-2000; 2000US-0191007P. 30-MAR-2000; 2000WO-US008439.
25-APR-2000; 2000US-0199397P. 22-MAY-2000; 2000WO-US014042.
05-JUN-2000; 2000US-0209832P.
Assignee: (GETH ) GENENTECH INC.
Inventors: Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
Cross reference: WPI; 2001-183260/18. N-PSDB; AAF92117.
Title: Eighty four nucleic acids encoding PRO polypeptides, useful in
molecular biology, including use as hybridization probes, and in
chromosome and gene mapping.
Patent format: Claim 12; Fig 120; 278pp; English.
Comment: The present sequence is a human PRO polypeptide (secreted and
transmembrane). The PRO protein, and PRO agonists, PRO antagonists
or anti-PRO antibodies are useful for preparation of a medicament
useful in the treatment of a condition which is responsive to the
PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO
protein may also be employed as molecular weight markers for
protein electrophoresis. The PRO coding sequence has applications
in molecular biology, including use as hybridisation probes, and in
chromosome and gene mapping
Database: GENESEQ patent database (v200420, 23-SEP-2004).

P_AAG89285 Human secreted protein, SEQ ID NO: 405 - Homo sapiens.
Length: 225 aa
Accession: P_AAG89285;
Species: Homo sapiens.
Keywords: Human; secreted protein; gene therapy; vaccine; treatment; diagnosis; GENSET; patent; GENESEQ patentdb.
Patent number: WO200142451-A2.
Publication date: 14-JUN-2001.
Filing date: 07-DEC-2000; 2000WO-IB001938.
Priority: 08-DEC-1999; 99US-0169629P. 06-MAR-2000; 2000US-0187470P.
Assignee: (GEST ) GENSET.
Inventors: Dumas Milne Edwards J, Bougueleret L, Jobert S;
Cross reference: WPI; 2001-367870/38. N-PSDB; AAH64888.
Title: Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.
Patent format: Claim 21; Page 885-886; 921pp; English.
Comment: The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy. The GENSET polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of GENSET polypeptide expression and activity. The present sequence is a GENSET polypeptide of the invention
Database: GENESEQ patent database (v200420, 23-SEP-2004).

P_AAM79556 Human protein SEQ ID NO 3202 - Homo sapiens.
Length: 231 aa
Accession: P_AAM79556;
Species: Homo sapiens.
Keywords: Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation; patent; GENESEQ patentdb.
Patent number: WO200157190-A2.
Publication date: 09-AUG-2001.
Filing date: 05-FEB-2001; 2001WO-US004098.
Priority: 03-FEB-2000; 2000US-00496914. 27-APR-2000; 2000US-00560875. 20-JUN-2000; 2000US-00598075. 19-JUL-2000; 2000US-00620325. 01-SEP-2000; 2000US-00654936. 15-SEP-2000; 2000US-00663561. 20-OCT-2000; 2000US-00693325. 30-NOV-2000; 2000US-00728422.
Assignee: (HYSE-) HYSEQ INC.
Inventors: Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y; Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW; Xue AJ,

Yang Y, Wejhrman T, Goodrich R;
Cross reference: WPI; 2001-476283/51. N-PSDB; AAK52689.
Title: Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.
Patent format: Claim 20; Page 283; 6221pp; English.
Comment: The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication
Database: GENESEQ patent database (v200420, 23-SEP-2004).

P_AAM78572 Human protein SEQ ID NO 1234 - Homo sapiens.
Length: 225 aa
Accession: P_AAM78572;
Species: Homo sapiens.
Keywords: Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation; patent; GENESEQ patentdb.
Patent number: WO200157190-A2.
Publication date: 09-AUG-2001.
Filing date: 05-FEB-2001; 2001WO-US004098.
Priority: 03-FEB-2000; 2000US-00496914. 27-APR-2000; 2000US-00560875. 20-JUN-2000; 2000US-00598075. 19-JUL-2000; 2000US-00620325. 01-SEP-2000; 2000US-00654936. 15-SEP-2000; 2000US-00663561. 20-OCT-2000; 2000US-00693325. 30-NOV-2000; 2000US-00728422.
Assignee: (HYSE-) HYSEQ INC.
Inventors: Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y; Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW; Xue AJ, Yang Y, Wejhrman T, Goodrich R;
Cross reference: WPI; 2001-476283/51. N-PSDB; AAK51705.
Title: Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.
Patent format: Claim 20; Page 3493-3494; 6221pp; English.
Comment: The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and

3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication
Database: GENESEQ patent database (v200420, 23-SEP-2004).

P_AAU29202 Human PRO polypeptide sequence #179 - Homo sapiens.
Length: 225 aa
Accession: P_AAU29202;
Species: Homo sapiens.
Keywords: PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha; blood; chondrocyte cell; cell proliferation; cell differentiation; colon; adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder; patent; GENESEQ patentdb.
Patent number: WO200168848-A2.
Publication date: 20-SEP-2001.
Filing date: 28-FEB-2001; 2001WO-US006520.
Priority: 01-MAR-2000; 2000WO-US005601. 02-MAR-2000; 2000WO-US005841. 03-MAR-2000; 2000US-0187202P.01-DEC-2000; 2000WO-US032678. 20-DEC-2000; 2000WO-US034956. plus 36 more dates.
Assignee: (GETH ) GENENTECH INC.
Inventors: Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL; Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
Cross reference: WPI; 2001-602746/68. N-PSDB; AAS46103.
Title: Novel nucleic acids encoding PRO polypeptides, used to diagnose the presence of tumors, such as prostate and breast tumors, in mammals and to screen for modulators of the compounds.
Patent format: Claim 11; Fig 358; 774pp; English.
Comment: Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention. The PRO polypeptides and their associated nucleic acids can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders
Database: GENESEQ patent database (v200420, 23-SEP-2004).

P_AAY99433 Human PRO1573 (UNQ779) amino acid sequence SEQ ID NO:328 - Homo sapiens.
Length: 225 aa
Accession: P_AAY99433;
Species: Homo sapiens.
Keywords: Human; PRO polypeptide; membrane bound protein; receptor; diagnosis; transmembrane; secretion; immunoadhesion; pharmaceutical; screening; patent; GENESEQ patentdb.
Patent number: WO200012708-A2.
Publication date: 09-MAR-2000.

Filing date: 01-SEP-1999; 99WO-US020111.
Priority: 01-SEP-1998; 98US-0098716P. 01-SEP-1998; 98US-0098749P. 01-SEP-1998; 98US-0098750P.18-NOV-1998; 98US-0108858P. 18-NOV-1998; 98US-0108904P. plus 119 more dates.
Assignee: (GETH ) GENENTECH INC.
Inventors: Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
Cross reference: WPI; 2000-237871/20. N-PSDB; AAA37115.
Title: New mammalian DNA sequences encoding transmembrane, receptor or secreted PRO polypeptides, useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions.
Patent format: Claim 12; Fig 188; 773pp; English.
Comment: AAA37022 to AAA37144 encode the new isolated human transmembrane, receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The transmembrane and receptor PRO proteins can be used for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions. The polypeptides and nucleotide sequences encoding then have various industrial applications, including uses as pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR primers and hybridisation probes used in the isolation of the PRO polypeptides from the present invention
Database: GENESEQ patent database (v200420, 23-SEP-2004).

BAA95567 CLDN8 /pid=BAA95567.1 - Homo sapiens
Length: 225 aa
Species: Homo sapiens (human)
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E., Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K., Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R., Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A., Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K., Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S., Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G., Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A., Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S., Hennig,S., Riesselmann,L., Dagand,E., Wehrmaeyer,S., Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and Yaspo,M.L., Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing Consortium: * RIKEN Genomic Sciences Center, Human Genome Research Group * Institute of Molecular Biotechnology, Genome Analysis * Keio University School of Medicine, Dept. of Molecular Biology * GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular Genetics (addresses see below) Title: Direct Submission
Locus: AP001707
Accession: AP001707 AL163252 BA000005
Cross-references: GI:7768788; BAA95567.1; AP001707_2
Database: GBTRANS

AAH20866 CLDN8 protein /pid=AAH20866.1 - Homo sapiens
Length: 225 aa
Species: Homo sapiens (human)
Strausberg,R., Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD

20892-2590, USA Title: Direct Submission
Gene: CLDN8
Locus: BC020866
Accession: BC020866
Cross-references: LocusID:9073; AAH20866.1; BC020866_1
Database: GBTRANS

CLD8_HUMAN Claudin-8 /pid=CAB60615.1 - homo sapiens
Length: 225 aa
Species: Homo sapiens (Human).
Accession: P56748; EMBL; AJ250711; CAB60615.1. EMBL; AY358707; AAQ89070.1. EMBL; AP001707; BAA95567.1. EMBL; BC020866; AAH20866.1. EMBL; BC058004; AAH58004.1. Genew; HGNC:2050; CLDN8. InterPro; IPR006187; Claudin. InterPro; IPR006188; Claudin_reg. InterPro; IPR004031; PMP22_Claudin. Pfam; PF00822; PMP22_Claudin; 1. PRINTS; PR01077; CLAUDIN. PROSITE; PS01346; CLAUDIN; 1.
Keen T.J., Inglehearn C.F., Submitted (nov-1999) to the Embl/genbank/ddbj databases. (ref. 1: sequence from n.a.)
Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J., Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P., Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S., Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J., Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J., Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A., Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.-H., Yansura D., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I., Godowski P., Gray A., Genome Res. 13, 2265-2270, 2003. MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003; (ref. 2: sequence from n.a. ) Title: "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.";
Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E., Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P., Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H., Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E., Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F., Lehrach H., Reinhardt R., Yaspo M.-L., Nature 405, 311-319, 2000. MEDLINE=20289799; PubMed=10830953; (ref. 3: sequence from n.a. ) Title: "The DNA sequence of human chromosome 21.";
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., Mcewan P.J., Mckernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez

A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A., Proc. Natl. Acad. Sci. U.S.A. 99, 16899-16903, 2002. MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899; (ref. 4: sequence from n.a.) Title: "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

Keywords: tight junction; transmembrane.
Taxid: tx:9606
Gene name: Name=CLDN8;
8-28/Domain: Transmem Potential.
82-102/Domain: Transmem Potential.
118-138/Domain: Transmem Potential.
167-187/Domain: Transmem Potential.
Database: Swissprot (SPROT), Release 40 (Jan 11, 2003)

P_AAG89273 Human secreted protein, SEQ ID NO: 393 - Homo sapiens.
Length: 225 aa
Accession: P_AAG89273;
Species: Homo sapiens.
Keywords: Human; secreted protein; gene therapy; vaccine; treatment; diagnosis; GENSET; patent; GENESEQ patentdb.
Patent number: WO200142451-A2.
Publication date: 14-JUN-2001.
Filing date: 07-DEC-2000; 2000WO-IB001938.
Priority: 08-DEC-1999; 99US-0169629P. 06-MAR-2000; 2000US-0187470P.
Assignee: (GEST ) GENSET.
Inventors: Dumas Milne Edwards J, Bougueleret L, Jobert S;
Cross reference: WPI; 2001-367870/38. N-PSDB; AAH64876.
Title: Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.
Patent format: Claim 21; Page 879; 921pp; English.
Comment: The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy. The GENSET polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of GENSET polypeptide expression and activity. The present sequence is a GENSET polypeptide of the invention
Database: GENESEQ patent database (v200420, 23-SEP-2004).

P_AAU00678 Human INTERCEPT 309 polypeptide - Homo sapiens.

Length: 215 aa
Accession: P_AAU00678;
Species: Homo sapiens.
Keywords: Human; INTERCEPT 309; T cell; heart; liver; pancreas; placenta; brain; skeletal muscle; kidney; spleen; lymph node; peripheral blood leukocyte; bone marrow; thymus tissue; abnormal blood coagulation; asthma; cancer; anaphylaxis; hepatitis; multiple sclerosis; coronary artery disease; malaria; atopic dermatitis; amyotrophic lateral sclerosis; meningitis; attention deficit disorder; Crohn's disease; gastroenteritis; goitre; hypoglycaemia; diabetes mellitus; endometriosis; pulmonary embolism; muscular dystrophy; immuno-competence; vertebrate; blood; serum; lung; patent; GENESEQ patentdb.
Patent number: WO200129088-A1.
Publication date: 26-APR-2001.
Filing date: 23-JUN-2000; 2000WO-US017386.
Priority: 19-OCT-1999; 99US-00420707.
Assignee: (MILL-) MILLENNIUM PHARM INC.
Inventors: Mackay CR, Myers PS, Kirst SJ, Fraser CC, Leiby KR;
Cross reference: WPI; 2001-308477/32. N-PSDB; AAS00668.
Title: New isolated nucleic acid molecule for diagnosis, prevention, and therapy of human and other animal disorder, or as modulating agent for regulating cellular processes.
Patent format: Claim 8; Fig 3A-3B; 263pp; English.
Comment: The sequence represents human INTERCEPT 309 polypeptide. This protein and similar others exhibit the ability to affect growth, proliferation, survival, differentiation, activity, morphology, or movement/migration of, e.g. T cells and cells of the heart, liver, pancreas, placenta, brain, lung, skeletal muscle, kidney, spleen, lymph node, peripheral blood leukocyte, bone marrow or thymus tissue. They can be used as modulating agents for regulating cellular processes, thus, the proteins and their associated nucleic acids can be used to prognosticate, prevent, diagnose, or treat disorders associated with physiological processes. These disorders include abnormal blood coagulation, asthma, anaphylaxis, hepatitis, multiple sclerosis, cancer, coronary artery disease, malaria, atopic dermatitis, amyotrophic lateral sclerosis, meningitis, attention deficit disorder, Crohn's disease, gastroenteritis, goitre, hypoglycaemia, diabetes mellitus, endometriosis, pulmonary embolism and muscular dystrophy. Antibodies to disorders such as these can be made by providing a polypeptide of the invention to an immuno-competent vertebrate and harvesting blood or serum from the vertebrate
1-24/Peptide
/note= Signal peptide/
25-215/Protein
/note= Mature human INTERCEPT 309/
25-71/Domain
/note= Extracellular domain/
72-92/Domain
/note= Transmembrane domain/
93-107/Domain
/note= Cytoplasmic domain/
108-131/Domain
/note= Transmembrane domain/
132-153/Domain
/note= Extracellular domain/

154-178/Domain
/note= Transmembrane domain/
179-215/Domain
/note= Cytoplasmic domain/
Database: GENESEQ patent database (v200420, 23-SEP-2004).

P_AAB93923 Human protein sequence SEQ ID NO:13910 - Homo sapiens.
Length: 233 aa
Accession: P_AAB93923;
Species: Homo sapiens.
Keywords: Human; primer; detection; diagnosis; antisense therapy; gene therapy; patent; GENESEQ patentdb.
Patent number: EP1074617-A2.
Publication date: 07-FEB-2001.
Filing date: 28-JUL-2000; 2000EP-00116126.
Priority: 29-JUL-1999; 99JP-00248036. 27-AUG-1999; 99JP-00300253. 11-JAN-2000; 2000JP-00118776. 02-MAY-2000; 2000JP-00183767. 09-JUN-2000; 2000JP-00241899.
Assignee: (HELI-) HELIX RES INST.
Inventors: Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J; Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
Cross reference: WPI; 2001-318749/34.
Title: Primer sets for synthesizing polynucleotides, particularly the 5602 full- length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
Patent format: Claim 8; SEQ ID NO 13910; 2537pp + Sequence Listing; English.
Comment: The present invention describes primer sets for synthesising 5602 full- length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
Database: GENESEQ patent database (v200420, 23-SEP-2004).

###
Tue Feb 5 15:17:50 2002 [BLASTN 2.2.1 [Jul-12-2001], NCBI]

/home/ruby/va/Molbio/carpenda/tempids/ss.DNA73735 (2010 bp)
/home/ruby/va/Molbio/carpenda/tempids/ss.DNA73735

| Sequences producing High-scoring Segment Pairs: | | Frame | Score | Match | Pct | E-val |
|---|---|---|---|---|---|---|
| 1 | P_AAF92117 Human PRO1573 cDNA. | + | 2010 | 2010 | 100 | 0.0 |
| 2 | P_AAS46103 Human DNA encoding PRO polypeptide seque | + | 2010 | 2010 | 100 | 0.0 |
| 3 | P_AAA37115 Human PRO1573 (UNQ779) cDNA sequence SEQ | + | 2010 | 2010 | 100 | 0.0 |
| 4 | AX092388 Sequence 119 from Patent WO0116318. DNA | + | 2010 | 2010 | 100 | 0.0 |
| 5 | AP000884 Homo sapiens genomic DNA, chromosome 21q | - | 2006 | 2009 | 100 | 0.0 |
| 6 | AP001707 Homo sapiens genomic DNA, chromosome 21q | - | 2006 | 2009 | 100 | 0.0 |
| 7 | HS21C052 Homo sapiens chromosome 21 segment HS21C | - | 2006 | 2009 | 100 | 0.0 |
| 8 | P_AAH64874 Human secreted protein cDNA, SEQ ID NO: | + | 1990 | 2005 | 100 | 0.0 |
| 9 | HSA250711 Homo sapiens CLDN8 gene for claudin-8. | + | 1927 | 1930 | 100 | 0.0 |
| 10 | P_AAS00668 Human INTERCEPT 309 cDNA. | + | 1879 | 1882 | 100 | 0.0 |
| 11 | P_AAH64876 Human secreted protein cDNA, SEQ ID NO: | + | 1811 | 1832 | 100 | 0.0 |
| 12 | P_AAH15597 Human cDNA sequence SEQ ID NO:13909. | + | 1794 | 1831 | 99 | 0.0 |
| 13 | AK022269 Homo sapiens cDNA FLJ12207 fis, clone MA | + | 1794 | 1831 | 99 | 0.0 |
| 14 | P_AAH64888 Human secreted protein cDNA, SEQ ID NO: | + | 1677 | 1692 | 100 | 0.0 |
| 15 | P_AAH64890 Human secreted protein cDNA, SEQ ID NO: | + | 1662 | 1689 | 99 | 0.0 |

GenBank (Release 143, aug 2004)

2010 100 0.0
P_AAF92117 Human PRO1573 cDNA. 010 bp, cDNA, PAT 15-MAY-2001
ACCESSION P_AAF92117
KEYWORDS GENESEQ; Human; PRO protein; mapping; patent; patentdb (v200420, 23-SEP-2004).
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2010)
AUTHORS Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A., Godowski,P.J. Grimaldi,C.J., Gurney,A.L., Watanabe,C.K., Wood,W.I.
TITLE Eighty four nucleic acids encoding PRO polypeptides, useful in molecular biology, including use as hybridization probes, and in chromosome and gene mapping.
JOURNAL Patent: WO200116318-A2; Filing Date: 24-AUG-2000; 2000WO-US023328; Publication Date: 08-MAR-2001; Priority: 01-SEP-1999; 99WO-US020111. 15-SEP-1999; 99WO-US021090. 07-DEC-1999; 99US-0169495P. 09-DEC-1999; 99US-0170262P. 11-JAN-2000; 2000US-0175481P. 18-FEB-2000; 2000WO-US004341. 18-FEB-2000; 2000WO-US004342. 22-FEB-2000; 2000WO-US004414. 01-MAR-2000; 2000WO-US005601. 03-MAR-2000; 2000US-0187202P. 21-MAR-2000; 2000US-0191007P. 30-MAR-2000; 2000WO-US008439. 25-APR-2000; 2000US-0199397P. 22-MAY-2000; 2000WO-US014042. 05-JUN-2000; 2000US-0209832P; Assignee: (GETH ) GENENTECH INC; Cross Reference: WPI; 2001-183260/18. P-PSDB; AAB87585; Patent Format: Claim 2; Fig 119; 278pp; English.
COMMENT The present sequence is the coding sequence for a human PRO polypeptide (secreted and transmembrane). The PRO protein, and PRO agonists, PRO antagonists or anti-PRO antibodies are useful for preparation of a medicament useful in the treatment of a condition which is responsive to the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO protein may also be employed as

molecular weight markers for protein electrophoresis. The PRO coding sequence has applications in molecular biology, including use as hybridisation probes, and in chromosome and gene mapping

FEATURES Location/Qualifiers

BASE COUNT 596 a 351 c 398 g 665 t

ORIGIN

2010 100 0.0

P_AAS46103 Human DNA encoding PRO polypeptide sequence #179. 010 bp, cDNA, PAT 18-DEC-2001

ACCESSION P_AAS46103

KEYWORDS GENESEQ; PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss; dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha; blood; chondrocyte cell; cell proliferation; cell differentiation; colon; adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder; PCR primer; patent; patentdb (v200420, 23-SEP-2004).

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 2010)

AUTHORS Baker,K.P., Chen,J., Desnoyers,L., Goddard,A., Godowski,P.J., Gurney,A.L. Pan,J., Smith,V., Watanabe,C.K., Wood,W.I., Zhang,Z.

TITLE Novel nucleic acids encoding PRO polypeptides, used to diagnose the presence of tumors, such as prostate and breast tumors, in mammals and to screen for modulators of the compounds.

JOURNAL Patent: WO200168848-A2; Filing Date: 28-FEB-2001; 2001WO-US006520; Publication Date: 20-SEP-2001; Priority: 01-MAR-2000; 2000WO-US005601. 02-MAR-2000; 2000WO-US005841. 03-MAR-2000; 2000US-0187202P. 06-MAR-2000; 2000US-0186968P. 14-MAR-2000; 2000US-0189320P. 14-MAR-2000; 2000US-0189328P. 15-MAR-2000; 2000WO-US006884. 21-MAR-2000; 2000US-0190828P. 21-MAR-2000; 2000US-0191007P. 21-MAR-2000; 2000US-0191048P. 21-MAR-2000; 2000US-0191314P. 28-MAR-2000; 2000US-0192655P. 29-MAR-2000; 2000US-0193032P. 29-MAR-2000; 2000US-0193053P. 30-MAR-2000; 2000WO-US008439. 04-APR-2000; 2000US-0194449P. 04-APR-2000; 2000US-0194647P. 11-APR-2000; 2000US-0195975P. 11-APR-2000; 2000US-0196000P. 11-APR-2000; 2000US-0196187P. 11-APR-2000; 2000US-0196690P. 11-APR-2000; 2000US-0196820P. 18-APR-2000; 2000US-0198121P. 18-APR-2000; 2000US-0198585P. 25-APR-2000; 2000US-0199397P. 25-APR-2000; 2000US-0199550P. 25-APR-2000; 2000US-0199654P. 03-MAY-2000; 2000US-0201516P. 17-MAY-2000; 2000WO-US013705. 22-MAY-2000; 2000WO-US014042. 30-MAY-2000; 2000WO-US014941. 02-JUN-2000; 2000WO-US015264. 05-JUN-2000; 2000US-0209832P. 28-JUL-2000; 2000WO-US020710. 22-AUG-2000; 2000US-00644848. 24-AUG-2000; 2000WO-US023328. 08-NOV-2000; 2000WO-US030952. 01-DEC-2000; 2000WO-US032678. 20-DEC-2000; 2000WO-US034956; Assignee: (GETH ) GENENTECH INC; Cross Reference: WPI; 2001-602746/68. P-PSDB; AAU29202; Patent Format: Claim 2; Fig 357; 774pp; English.

COMMENT Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR primers for PRO polypeptides of the invention. The sequences of the invention can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats and rabbits but are

preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders

FEATURES Location/Qualifiers

BASE COUNT 596 a 351 c 398 g 665 t

ORIGIN

2010 100 0.0

P_AAA37115 Human PRO1573 (UNQ779) cDNA sequence SEQ ID NO:327. 010 bp, cDNA, PAT 08-AUG-2000

ACCESSION P_AAA37115

KEYWORDS GENESEQ; Human; PRO polypeptide; membrane bound protein; receptor; diagnosis; transmembrane; secretion; immunoadhesion; pharmaceutical; screening; patent; patentdb (v200420, 23-SEP-2004).

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 2010)

AUTHORS Baker,K., Goddard,A., Gurney,A.L., Smith,V., Watanabe,C.K., Wood,W.I.

TITLE New mammalian DNA sequences encoding transmembrane, receptor or secreted PRO polypeptides, useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions.

JOURNAL Patent: WO200012708-A2; Filing Date: 01-SEP-1999; 99WO-US020111; Publication Date: 09-MAR-2000; Priority: 01-SEP-1998; 98US-0098716P. 01-SEP-1998; 98US-0098749P. 01-SEP-1998; 98US-0098750P. 02-SEP-1998; 98US-0098803P. 02-SEP-1998; 98US-0098821P. 02-SEP-1998; 98US-0098843P. 09-SEP-1998; 98US-0099536P. 09-SEP-1998; 98US-0099596P. 09-SEP-1998; 98US-0099598P. 09-SEP-1998; 98US-0099602P. 09-SEP-1998; 98US-0099642P. 10-SEP-1998; 98US-0099741P. 10-SEP-1998; 98US-0099754P. 10-SEP-1998; 98US-0099763P. 10-SEP-1998; 98US-0099792P. 10-SEP-1998; 98US-0099808P. 10-SEP-1998; 98US-0099812P. 10-SEP-1998; 98US-0099815P. 10-SEP-1998; 98US-0099816P. 15-SEP-1998; 98US-0100385P. 15-SEP-1998; 98US-0100388P. 15-SEP-1998; 98US-0100390P. 16-SEP-1998; 98US-0100584P. 16-SEP-1998; 98US-0100627P. 16-SEP-1998; 98US-0100661P. 16-SEP-1998; 98US-0100662P. 16-SEP-1998; 98US-0100664P. 17-SEP-1998; 98US-0100683P. 17-SEP-1998; 98US-0100684P. 17-SEP-1998; 98US-0100710P. 17-SEP-1998; 98US-0100711P. 17-SEP-1998; 98US-0100919P. 17-SEP-1998; 98US-0100930P. 18-SEP-1998; 98US-0100848P. 18-SEP-1998; 98US-0100849P. 18-SEP-1998; 98US-0101014P. 18-SEP-1998; 98US-0101068P. 18-SEP-1998; 98US-0101071P. 22-SEP-1998; 98US-0101279P. 23-SEP-1998; 98US-0101471P. 23-SEP-1998; 98US-0101472P. 23-SEP-1998; 98US-0101474P. 23-SEP-1998; 98US-0101475P. 23-SEP-1998; 98US-0101476P. 23-SEP-1998; 98US-0101477P. 23-SEP-1998; 98US-0101479P. 24-SEP-1998; 98US-0101738P. 24-SEP-1998; 98US-0101741P. 24-SEP-1998; 98US-0101743P. 24-SEP-1998; 98US-0101915P. 24-SEP-1998;

98US-0101916P. 29-SEP-1998; 98US-0102207P. 29-SEP-1998;
98US-0102240P. 29-SEP-1998; 98US-0102307P. 29-SEP-1998;
98US-0102330P. 29-SEP-1998; 98US-0102331P. 30-SEP-1998;
98US-0102484P. 30-SEP-1998; 98US-0102487P. 30-SEP-1998;
98US-0102570P. 30-SEP-1998; 98US-0102571P. 01-OCT-1998;
98US-0102684P. 01-OCT-1998; 98US-0102687P. 02-OCT-1998;
98US-0102965P. 06-OCT-1998; 98US-0103258P. 06-OCT-1998;
98US-0103449P. 07-OCT-1998; 98US-0103314P. 07-OCT-1998;
98US-0103315P. 07-OCT-1998; 98US-0103328P. 07-OCT-1998;
98US-0103395P. 07-OCT-1998; 98US-0103396P. 07-OCT-1998;
98US-0103401P. 08-OCT-1998; 98US-0103633P. 08-OCT-1998;
98US-0103678P. 08-OCT-1998; 98US-0103679P. 08-OCT-1998;
98US-0103711P. 14-OCT-1998; 98US-0104257P. 20-OCT-1998;
98US-0104987P. 20-OCT-1998; 98US-0105000P. 20-OCT-1998;
98US-0105002P. 21-OCT-1998; 98US-0105104P. 22-OCT-1998;
98US-0105169P. 22-OCT-1998; 98US-0105266P. 26-OCT-1998;
98US-0105693P. 26-OCT-1998; 98US-0105694P. 27-OCT-1998;
98US-0105807P. 27-OCT-1998; 98US-0105881P. 27-OCT-1998;
98US-0105882P. 27-OCT-1998; 98US-0106062P. 28-OCT-1998;
98US-0106023P. 28-OCT-1998; 98US-0106029P. 28-OCT-1998;
98US-0106030P. 28-OCT-1998; 98US-0106032P. 28-OCT-1998;
98US-0106033P. 28-OCT-1998; 98US-0106178P. 29-OCT-1998;
98US-0106248P. 29-OCT-1998; 98US-0106384P. 29-OCT-1998;
98US-0108500P. 30-OCT-1998; 98US-0106464P. 03-NOV-1998;
98US-0106856P. 03-NOV-1998; 98US-0106902P. 03-NOV-1998;
98US-0106905P. 03-NOV-1998; 98US-0106919P. 03-NOV-1998;
98US-0106932P. 03-NOV-1998; 98US-0106934P. 10-NOV-1998;
98US-0107783P. 17-NOV-1998; 98US-0108775P. 17-NOV-1998;
98US-0108779P. 17-NOV-1998; 98US-0108787P. 17-NOV-1998;
98US-0108788P. 17-NOV-1998; 98US-0108801P. 17-NOV-1998;
98US-0108802P. 17-NOV-1998; 98US-0108806P. 17-NOV-1998;
98US-0108807P. 17-NOV-1998; 98US-0108867P. 17-NOV-1998;
98US-0108925P. 18-NOV-1998; 98US-0108848P. 18-NOV-1998;
98US-0108849P. 18-NOV-1998; 98US-0108850P. 18-NOV-1998;
98US-0108851P. 18-NOV-1998; 98US-0108852P. 18-NOV-1998;
98US-0108858P. 18-NOV-1998; 98US-0108904P; Assignee: (GETH )
GENENTECH INC; Cross Reference: WPI; 2000-237871/20. P-PSDB;
AAY99433; Patent Format: Claim 2; Fig 187; 773pp; English.

COMMENT AAA37022 to AAA37144 encode the new isolated human transmembrane, receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The transmembrane and receptor PRO proteins can be used for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions. The polypeptides and nucleotide sequences encoding then have various industrial applications, including uses as pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR primers and hybridisation probes used in the isolation of the PRO polypeptides from the present invention

FEATURES Location/Qualifiers

BASE COUNT 596 a 351 c 398 g 665 t

ORIGIN

2010 100 0.0

AX092388 Sequence 119 from Patent WO0116318. 2010 bp, DNA, linear, PAT 21-MAR-2001

ACCESSION AX092388

VERSION AX092388.1 GI:13444508

```
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A.,
            Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and
            Wood,W.I.
  TITLE     Secreted and transmembrane polypeptides and nucleic acids encoding
            the same
  JOURNAL   Patent: WO 0116318-A 119 08-MAR-2001;
            Genentech, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..2010
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
BASE COUNT
ORIGIN

2009 100   0.0
AP000884    Homo sapiens genomic DNA, chromosome 21q22.1, clone:B463J19,
            SOD-AML region, complete sequence. 137955 bp,
            DNA, linear, PRI 16-DEC-1999
ACCESSION   AP000884
VERSION     AP000884.1  GI:6580113
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 137955)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens 137,955 genomic DNA of 21q22.1
  JOURNAL   Published Only in DataBase (1999)
REFERENCE   2  (bases 1 to 137955)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-DEC-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
FEATURES             Location/Qualifiers
     source          1..137955
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="B463J19"
BASE COUNT
ORIGIN
```

2009 100  0.0
AP001707  Homo sapiens genomic DNA, chromosome 21q, section 51/105. 340000 bp, DNA, linear, PRI 21-MAY-2003
ACCESSION  AP001707 AL163252 BA000005
VERSION  AP001707.1  GI:7768786
KEYWORDS  .
SOURCE  Homo sapiens (human)
  ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
  AUTHORS  Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E., Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K., Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R., Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A., Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K., Asakawa,S., Shintani,A., Sasaki,K., Nagamine,K., Mitsuyama,S., Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G., Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A., Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S., Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and Yaspo,M.L.
  TITLE  The DNA sequence of human chromosome 21
  JOURNAL  Nature 405 (6784), 311-319 (2000)
  MEDLINE  20289799
   PUBMED  10830953
REFERENCE  2  (bases 1 to 340000)
  AUTHORS  Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E., Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K., Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R., Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A., Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K., Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S., Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G., Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A., Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S., Hennig,S., Riesselmann,L., Dagand,E., Wehrmaeyer,S., Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and Yaspo,M.L.
  TITLE  Direct Submission
  JOURNAL  Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing Consortium: * RIKEN Genomic Sciences Center, Human Genome Research Group * Institute of Molecular Biotechnology, Genome Analysis * Keio University School of Medicine, Dept. of Molecular Biology * GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular Genetics (addresses see below)
COMMENT  On Jan 16, 2002 this sequence version replaced gi:7717317.
The chromosome 21 mapping and sequencing consortium consisting of
* RIKEN Genomic Sciences Center, Human Genome Research Group, *
Sagamihara 228-8555, Japan,
* e.mail: hattori@gsc.riken.go.jp
* URL: http://hgp.gsc.riken.go.jp/
and

```
            * Institute of Molecular Biotechnology, Genome Analysis, *
            Beutenbergstrasse 11, D-07745 Jena, Germany,
            * e.mail: gscj-submit@genome.imb-jena.de
            * URL: http://genome.imb-jena.de/
            and
            * Keio University School of Medicine, Molecular Biology, * Tokyo
            160-8582, Japan,
            * e.mail: nshimizu@dmb-med.keio.ac.jp
            * URL: http://www.dmb.med.keio.ac.jp/
            and
            * GBF, Dept. of Genome Analysis,
            * Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
            info.genome@gbf.de
            * URL: http://genome.gbf.de/
            and
            * Max-Planck Institute for Molecular Genetics,
            * Ihnestrasse 73, D-14195 Berlin, Germany,
            * e.mail: info-chr21@molgen.mpg.de
            * URL: http://chr21.rz-berlin.mpg.de/
            AL163252: Submitted (10-Apr-2000).
FEATURES             Location/Qualifiers
     source          1..340000
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
     source          <1..148
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="R355D16, 5' partial"
                     /clone_lib="RPCI-11 BAC library"
                     /note="Accession No. AP000236"
     source          <1..48622
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="R347N6, 5' partial"
                     /clone_lib="RPCI-11 BAC library"
                     /note="Accession No. AP000390"
     source          <1..57177
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="P681K20, 5' partial"
                     /clone_lib="RPCI1,3-5 PAC library"
                     /note="Accession No. AF191069"
     source          26572..135558
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
```

```
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="P131P23"
                     /clone_lib="RPCI1,3-5 PAC library"
                     /note="Accession No. AF190641"
     source          48469..158332
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="P209K9"
                     /clone_lib="RPCI1,3-5 PAC library"
                     /note="Accession No. AF096876"
     source          156690..294644
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="R463J19"
                     /clone_lib="RPCI-11 BAC library"
                     /note="Accession No. AP000884"
     source          292534..}340000
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="R680H4, 3' partial"
                     /clone_lib="RPCI-11 BAC library"
                     /note="Accession No. AP000885"
     repeat_region   1610..1766
                     /note="MER50"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region   1857..2051
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   2054..2392
                     /note="MER50"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region   3327..3518
                     /note="MER102"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   complement(3582..3694)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   4580..4614
                     /note="(TTC)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   5743..6052
```

```
                     /note="AluSq"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(6134..6214)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   complement(6299..6651)
                     /note="MLT1A1"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   7531..7839
                     /note="AluSc"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   9145..9173
                     /note="(A)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   complement(10285..10467)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   10468..10753
                     /note="AluJb"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   10757..10785
                     /note="(TAAAA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   complement(10786..10847)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   12268..13078
                     /note="L1M4"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   13293..13411
                     /note="L1MEc"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   13442..13591
                     /note="L1MEc"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   13704..13750
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   complement(15714..16207)
                     /note="HERVL"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region   complement(16208..16352)
                     /note="AluSg/x"
```

```
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     complement(16355..16896)
                     /note="HERVL"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region     complement(16897..17168)
                     /note="AluSg"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     complement(17169..17290)
                     /note="HERVL"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region     complement(17288..17388)
                     /note="HERVL"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region     complement(17405..17752)
                     /note="HERVL"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region     17776..17856
                     /note="(TA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region     17875..17919
                     /note="(TATG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region     17948..18035
                     /note="L1M4"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     18040..18580
                     /note="L1MB8"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     complement(18692..18805)
                     /note="HERVL"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region     complement(19287..19701)
                     /note="MLT2H"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region     20399..20763
                     /note="MLT1A1"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region     complement(20896..21252)
                     /note="MER69"
                     /rpt_family="DNA/AcHobo"
                     /rpt_type=DISPERSED
     repeat_region     complement(21253..21393)
                     /note="FLAM_C"
                     /rpt_family="SINE/Alu"
```

```
                     /rpt_type=DISPERSED
     repeat_region     complement(21398..21995)
                     /note="MER69"
                     /rpt_family="DNA/AcHobo"
                     /rpt_type=DISPERSED
     repeat_region     22187..22495
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     22496..22515
                     /note="(GAAAA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region     23349..23361
                     /note="(TAA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region     23362..23559
                     /note="AluSg/x"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     23560..23581
                     /note="(TAA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     STS               23648..23791
                     /standard_name="D21S1846"
                     /note="Accession No. D42163
                     HY21-285aR"
     repeat_region     complement(23846..24948)
                     /note="MER69"
                     /rpt_family="DNA/AcHobo"
                     /rpt_type=DISPERSED
     repeat_region     complement(24949..25275)
                     /note="MER61A"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region     complement(25279..25594)
                     /note="MER69"
                     /rpt_family="DNA/AcHobo"
                     /rpt_type=DISPERSED
     repeat_region     complement(28100..28223)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region     complement(28403..28727)
                     /note="MER33"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region     complement(28823..28942)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region     complement(29446..29820)
                     /note="MLT1A2"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
```

```
repeat_region   29905..30078
                /note="MER5B"
                /rpt_family="DNA/MER1_type"
                /rpt_type=DISPERSED
repeat_region   30446..30598
                /note="MER5A"
                /rpt_family="DNA/MER1_type"
                /rpt_type=DISPERSED
repeat_region   complement(31211..31515)
                /note="AluSx"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   31729..31746
                /note="(TG)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   31747..31783
                /note="(TG)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   31813..31912
                /note="(CATATA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   complement(32487..32636)
                /note="LTR27"
                /rpt_family="LTR/MER4-group"
                /rpt_type=DISPERSED
repeat_region   complement(32694..33198)
                /note="LTR27"
                /rpt_family="LTR/MER4-group"
                /rpt_type=DISPERSED
repeat_region   33408..33579
                /note="MLT1D"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   34013..34091
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   34093..34277
                /note="(TCTCC)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   34611..34876
                /note="L1MC1"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(35441..35584)
                /note="MER58B"
                /rpt_family="DNA/MER1_type"
                /rpt_type=DISPERSED
repeat_region   complement(35585..35767)
                /note="AluY"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   complement(35768..35824)
```

```
                     /note="MER58B"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region     36150..36200
                     /note="(TA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region     complement(37283..37594)
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     38551..38618
                     /note="(TTC)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region     39174..39302
                     /note="L1MD3"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     complement(39345..39510)
                     /note="L1M4"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     39512..39717
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     39718..40146
                     /note="LTR7"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region     40148..40244
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     complement(40247..40461)
                     /note="L1M4"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     complement(40647..40885)
                     /note="L1MEc"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     complement(40946..42900)
                     /note="L1M4"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     44069..44189
                     /note="AluJb"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     44190..44212
                     /note="(CAT)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region     44215..44256
                     /note="BC200"
```

```
                     /rpt_family="scRNA"
                     /rpt_type=DISPERSED
     repeat_region   45825..45891
                     /note="(TG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   47490..48123
                     /note="MER41B"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region   complement(48305..48393)
                     /note="MER5A"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   complement(48394..48702)
                     /note="AluSg"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(48703..48766)
                     /note="MER5A"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   48800..49452
                     /note="PABL_A"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region   49471..49621
                     /note="MER5B"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   complement(49915..50054)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   50849..51045
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   complement(52528..52769)
                     /note="Charlie1a"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   complement(52777..53603)
                     /note="L1PA13"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(53604..53831)
                     /note="Charlie1a"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   53832..54122
                     /note="AluSg"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(54123..55116)
                     /note="Charlie1a"
                     /rpt_family="DNA/MER1_type"
```

```
                     /rpt_type=DISPERSED
     repeat_region   complement(55804..55958)
                     /note="L1MA2"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(56025..56407)
                     /note="THE1C"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   56596..56615
                     /note="(TTG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   56906..57196
                     /note="AluSg"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   57276..57585
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(57919..58257)
                     /note="L1PA14"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(59538..59846)
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   60284..60561
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   60878..61021
                     /note="MER5A"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   61022..61304
                     /note="AluSp"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   61305..61359
                     /note="(TAAAA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   61361..61378
                     /note="MER5A"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   complement(62037..62224)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   62593..62642
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
```

```
repeat_region   64995..65340
                /note="L1ME3A"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(65614..66155)
                /note="L1M4"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(66210..66282)
                /note="L1ME"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   66285..66307
                /note="(TTTTA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   complement(66308..66463)
                /note="AluSg/x"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   complement(66465..67226)
                /note="L1ME"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   68693..68831
                /note="MER104"
                /rpt_family="DNA/MER1_type"
                /rpt_type=DISPERSED
repeat_region   complement(70650..70847)
                /note="AluY"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   complement(72537..72899)
                /note="MLT1A1"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   73742..74087
                /note="THE1B"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   75666..76000
                /note="THE1B"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   76064..76534
                /note="MLT1D"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   complement(77488..77640)
                /note="AluJo"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   complement(78204..78675)
                /note="MLT1H"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   79704..79724
```

```
                     /note="(TG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   complement(79859..80105)
                     /note="AluSg"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   81177..81331
                     /note="MER53"
                     /rpt_family="DNA"
                     /rpt_type=DISPERSED
     repeat_region   82207..82702
                     /note="MER66B"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region   84291..84769
                     /note="L1MEc"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   84895..86159
                     /note="L1MEc"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(86173..86603)
                     /note="MLT1C"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   87059..87083
                     /note="(TG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   88474..88507
                     /note="(TC)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   88508..88527
                     /note="(CA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   88596..89068
                     /note="MLT1F1"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   89598..89648
                     /note="(CATA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   91581..92953
                     /note="L1PA8"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   93138..93279
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   complement(94478..94637)
                     /note="MER5A"
```

```
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   94800..94858
                     /note="L1"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   94855..95590
                     /note="L1PA16"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(95618..95915)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     STS             96719..97040
                     /standard_name="D21S93"
                     /db_xref="GDB:192003"
     repeat_region   97850..97997
                     /note="MER5A"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   98340..98603
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   100248..100273
                     /note="(TCTA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   101340..101376
                     /note="(CA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   101803..101903
                     /note="MADE1"
                     /rpt_family="DNA/Mariner"
                     /rpt_type=DISPERSED
     repeat_region   complement(104577..104851)
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   104868..104994
                     /note="AluJo"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(104995..105295)
                     /note="AluSg"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   105296..105496
                     /note="AluJo"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   106248..106271
                     /note="(CA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
```

```
repeat_region   complement(106998..107232)
                /note="L1M4"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(107236..108368)
                /note="L1MB8"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   110671..110915
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   complement(111090..111456)
                /note="MLT1B"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   116374..116397
                /note="(GGAA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   117041..117786
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   117922..117945
                /note="(TTTTG)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   complement(117949..118210)
                /note="AluSc"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   118329..118723
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   complement(119645..119734)
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   complement(122484..122818)
                /note="L1MA1"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(122910..123004)
                /note="MER112"
                /rpt_family="DNA/MER1_type"
                /rpt_type=DISPERSED
repeat_region   123306..123609
                /note="AluSg"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   123733..123874
                /note="MER103"
                /rpt_family="DNA"
                /rpt_type=DISPERSED
repeat_region   complement(124583..124659)
```

```
                     /note="MER5A"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
repeat_region        complement(125311..125442)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
repeat_region        126521..126969
                     /note="L1M4"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
repeat_region        complement(127168..127470)
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
repeat_region        127973..128249
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
repeat_region        128250..128276
                     /note="(CAAAA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
repeat_region        complement(128361..129106)
                     /note="LTR1B"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
repeat_region        129465..129551
                     /note="MER5A"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
repeat_region        complement(130876..130938)
                     /note="L1PB2"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
repeat_region        complement(130940..131303)
                     /note="THE1C"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
repeat_region        complement(131311..132911)
                     /note="THE1C-internal"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
repeat_region        complement(132914..133274)
                     /note="THE1C"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
repeat_region        complement(133282..134013)
                     /note="L1PB2"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
repeat_region        134012..134613
                     /note="L1PB2"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
STS                  135457..135593
                     /standard_name="D21S1695"
```

```
                     /note="232C7aL
                     Accession No. D38269"
     repeat_region     complement(136167..136278)
                     /note="MLT1I"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region     complement(138259..139869)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region     complement(139966..140245)
                     /note="AluJo"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     complement(140496..141823)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region     142802..142991
                     /note="MER41A"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region     142992..143282
                     /note="AluSc"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     143283..143655
                     /note="MER41A"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region     143854..144372
                     /note="MER4B"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region     complement(144887..145196)
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     complement(145960..146456)
                     /note="LTR19A"
                     /rpt_family="LTR"
                     /rpt_type=DISPERSED
     repeat_region     complement(148453..148757)
                     /note="AluJo"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     complement(149220..149721)
                     /note="MLT1F"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region     complement(149774..149906)
                     /note="FLAM_C"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     complement(150819..151269)
                     /note="L2"
                     /rpt_family="LINE/L2"
```

```
                     /rpt_type=DISPERSED
     repeat_region   complement(151444..151618)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   151947..152069
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   complement(152112..152205)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   152215..152383
                     /note="(TCTA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   complement(153794..153962)
                     /note="MER5A"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   complement(154253..154624)
                     /note="L1M2"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(155818..156174)
                     /note="THE1B"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   complement(156354..156451)
                     /note="AluYb8"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(156453..156484)
                     /note="MLT1H"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   157012..157102
                     /note="MER94"
                     /rpt_family="DNA"
                     /rpt_type=DISPERSED
     repeat_region   complement(157361..157467)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   157517..157596
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   157618..157705
                     /note="(TATATG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   complement(159423..159610)
                     /note="LTR33A"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
```

```
repeat_region   159865..159943
                /note="MER97c"
                /rpt_family="DNA/MER1_type?"
                /rpt_type=DISPERSED
repeat_region   160006..160108
                /note="MER97a"
                /rpt_family="DNA/MER1_type?"
                /rpt_type=DISPERSED
repeat_region   160309..160470
                /note="(TCTA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   complement(160471..160878)
                /note="MLT1B"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   complement(160891..161131)
                /note="MER97a"
                /rpt_family="DNA/MER1_type?"
                /rpt_type=DISPERSED
repeat_region   161890..162041
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   complement(162554..162614)
                /note="MER5B"
                /rpt_family="DNA/MER1_type"
                /rpt_type=DISPERSED
gene            complement(164968..165642)
                /gene="CLDN17"
mRNA            complement(164968..165642)
                /gene="CLDN17"
                /note="human gene for claudin-17, Accession No. AJ250712"
CDS             complement(164968..165642)
                /gene="CLDN17"
                /note="human gene for claudin-17, Accession No. AJ250712"
                /codon_start=1
                /protein_id="BAA95566.1"
                /db_xref="GI:7768787"
exon            complement(164968..165642)
                /gene="CLDN17"
                /number=1
repeat_region   167617..167889
                /note="AluJb"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   168032..168186
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   168195..168527
                /note="L1PA16"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(168528..168606)
                /note="AluSc"
                /rpt_family="SINE/Alu"
```

```
                     /rpt_type=DISPERSED
     repeat_region   168648..168679
                     /note="(TAAAA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   168928..169108
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   169447..169805
                     /note="LTR16A"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region   169861..170088
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   complement(170145..170646)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   170647..170948
                     /note="AluSg"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(170949..171405)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   171759..171785
                     /note="(TTTG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   complement(171786..172071)
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   172208..172233
                     /note="(CA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   172349..172598
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   172646..172710
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   173181..173215
                     /note="(T)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   173639..173767
                     /note="MER5B"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
```

```
repeat_region   173799..173859
                /note="MER5B"
                /rpt_family="DNA/MER1_type"
                /rpt_type=DISPERSED
repeat_region   173877..174053
                /note="MER5B"
                /rpt_family="DNA/MER1_type"
                /rpt_type=DISPERSED
repeat_region   complement(175682..175736)
                /note="L1MC3"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(175737..176033)
                /note="AluSx"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   complement(176034..176617)
                /note="L1MC3"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(176621..176802)
                /note="L1MC3"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(176803..177103)
                /note="AluJb"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   complement(177104..177219)
                /note="L1MC3"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   178433..178794
                /note="MLT1D"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   178805..179028
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   complement(179612..179652)
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   179802..179976
                /note="MER20"
                /rpt_family="DNA/MER1_type"
                /rpt_type=DISPERSED
repeat_region   complement(179979..180204)
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   complement(181241..181414)
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   181545..181591
```

```
                     /note="MER20B"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   181592..181632
                     /note="(TG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   181633..182293
                     /note="MER20B"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   complement(182376..182500)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   182510..182779
                     /note="LTR37B"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region   182818..182897
                     /note="LTR37B"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region   183294..183587
                     /note="AluSq"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   184702..184877
                     /note="HAL1"
                     /rpt_family="LINE/Other"
                     /rpt_type=DISPERSED
     repeat_region   185141..185593
                     /note="MLT1D"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   186761..186852
                     /note="L1PA8"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   188430..188561
                     /note="L1PA15"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(188895..189066)
                     /note="MLT1J"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   complement(189265..189339)
                     /note="MLT1J"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   189795..190352
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   complement(190366..190506)
                     /note="MER5B"
```

```
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region     complement(190558..190629)
                     /note="MER4A2"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region     complement(190625..190745)
                     /note="MER4A"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region     complement(190777..191146)
                     /note="MLT1D"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region     complement(191227..191346)
                     /note="MLT1D"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region     191353..192586
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region     complement(192679..193267)
                     /note="MER89"
                     /rpt_family="LTR"
                     /rpt_type=DISPERSED
     repeat_region     193396..193432
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region     complement(194228..194449)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region     194909..194951
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region     195010..195029
                     /note="(TG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region     196293..196403
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region     197062..197425
                     /note="L1MA5"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     complement(197429..197480)
                     /note="L1P"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     197479..198709
                     /note="L1P"
                     /rpt_family="LINE/L1"
```

```
                     /rpt_type=DISPERSED
     repeat_region   198710..198984
                     /note="L1PA7"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   198999..199018
                     /note="(TAAAA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   complement(199526..200246)
                     /note="L1PA7"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(200247..201488)
                     /note="L1P"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(201590..201653)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   202619..202926
                     /note="AluJo"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(203131..203264)
                     /note="FLAM_C"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   203305..203433
                     /note="FLAM_A"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   203607..203646
                     /note="(GA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   complement(203963..204053)
                     /note="MER67C"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region   complement(204054..204361)
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(204362..204813)
                     /note="MER67C"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region   complement(204820..205121)
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   205193..205219
                     /note="(TTG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
```

```
repeat_region   complement(205225..205419)
                /note="L1PA4"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   205492..205569
                /note="MLT1G1"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   205751..205854
                /note="MLT1H"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   complement(205868..206401)
                /note="LTR26E"
                /rpt_family="LTR/Retroviral"
                /rpt_type=DISPERSED
repeat_region   207212..207472
                /note="AluJo"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   207480..207656
                /note="(GAAA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   209242..209271
                /note="(TA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   209299..209466
                /note="(TA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   209490..209519
                /note="(TA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   209588..209619
                /note="(TA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   complement(210077..210309)
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   complement(211664..211851)
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   212184..212355
                /note="AluSg/x"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
gene            complement(213042..214972)
                /gene="CLDN8"
mRNA            complement(213042..214972)
                /gene="CLDN8"
                /note="human gene for claudin-8, Accession No. AJ250711"
```

```
exon            complement(213042..214972)
                /gene="CLDN8"
                /number=1
5'UTR           complement(213042..214272)
                /gene="CLDN8"
polyA_signal    complement(213055..213060)
                /gene="CLDN8"
CDS             complement(214273..214950)
                /gene="CLDN8"
                /note="human gene for claudin-8, Accession No. AJ250711"
                /codon_start=1
                /protein_id="BAA95567.1"
                /db_xref="GI:7768788"
5'UTR           complement(214951..214972)
                /gene="CLDN8"
repeat_region   216645..216952
                /note="AluY"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   complement(218710..219021)
                /note="AluSx"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   complement(219103..219477)
                /note="MLT1A2"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   complement(219991..220158)
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   complement(221295..221472)
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   221763..221790
                /note="(CA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   complement(222287..222349)
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   223032..223125
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   223233..223386
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   complement(225324..225394)
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   225395..225716
                /note="AluSx"
```

```
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     complement(225717..225791)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region     225866..226083
                     /note="MER102"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region     227539..227832
                     /note="AluJb"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     227864..227891
                     /note="(A)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region     228396..228475
                     /note="(TA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region     230433..230722
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     230885..231744
                     /note="L1PA16"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     231745..232185
                     /note="L1PA16"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     complement(233642..233846)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region     complement(234030..234188)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region     complement(235129..235292)
                     /note="LTR23"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region     complement(235293..235575)
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     complement(235576..235617)
                     /note="LTR23"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region     complement(235618..235925)
                     /note="AluSg"
                     /rpt_family="SINE/Alu"
```

```
                     /rpt_type=DISPERSED
     repeat_region   complement(235926..236188)
                     /note="LTR23"
                     /rpt_family="LTR/MER4-group"
                     /rpt_type=DISPERSED
     repeat_region   236576..236642
                     /note="(TTTTC)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   complement(236644..236951)
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   239604..239738
                     /note="LTR47A"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region   complement(239739..239890)
                     /note="FLAM_C"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   239891..240156
                     /note="LTR47A"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region   240273..240301
                     /note="(TTTG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   complement(240307..240584)
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(240682..240800)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   complement(242367..242423)
                     /note="Alu"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   242425..242635
                     /note="AluSg/x"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   243792..244003
                     /note="L1ME"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(244757..244991)
                     /note="AluJb"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   244994..245575
                     /note="L1ME3A"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
```

```
repeat_region   245576..245881
                /note="AluJo"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   245882..245960
                /note="L1ME3A"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   246003..246717
                /note="L1ME3A"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   246750..246919
                /note="L1ME3A"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   246920..246995
                /note="FRAM/FAM"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   247035..247344
                /note="L1ME3A"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(247345..247423)
                /note="MADE1"
                /rpt_family="DNA/Mariner"
                /rpt_type=DISPERSED
repeat_region   247424..247561
                /note="L1ME3A"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   247909..248045
                /note="MSTA"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   248103..248149
                /note="MSTA"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   249668..250027
                /note="THE1B"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   complement(250448..250819)
                /note="MLT2FB"
                /rpt_family="LTR/Retroviral"
                /rpt_type=DISPERSED
repeat_region   251644..251799
                /note="(TA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   251918..252016
                /note="(TATATG)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   252054..252503
```

```
                     /note="L1MA3"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   252504..252549
                     /note="MADE1"
                     /rpt_family="DNA/Mariner"
                     /rpt_type=DISPERSED
     repeat_region   252550..253015
                     /note="L1MA3"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   253016..253314
                     /note="AluSq"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   253315..253431
                     /note="L1MA3"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   253795..253870
                     /note="MLT1L"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   complement(254749..255256)
                     /note="MLT2CB"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region   complement(255877..256168)
                     /note="AluJo"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(256689..256767)
                     /note="AluJb"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   257615..257663
                     /note="(TG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   257678..257803
                     /note="MER20"
                     /rpt_family="DNA/MER1_type"
                     /rpt_type=DISPERSED
     repeat_region   257864..258022
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   259701..260270
                     /note="L1PA13"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(261342..261638)
                     /note="AluSq"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   261745..262041
                     /note="AluSx"
```

```
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   263606..263658
                     /note="MER46A"
                     /rpt_family="DNA/MER2_type"
                     /rpt_type=DISPERSED
     repeat_region   complement(263659..263748)
                     /note="L1PA9"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   263749..263929
                     /note="MER46A"
                     /rpt_family="DNA/MER2_type"
                     /rpt_type=DISPERSED
     repeat_region   complement(263948..264193)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   264314..264542
                     /note="L1MA2"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   264678..265115
                     /note="LTR3"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region   complement(265834..265929)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   complement(265977..266095)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   complement(266377..266741)
                     /note="MLT1A1"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   complement(267097..267399)
                     /note="AluSq"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   267891..268019
                     /note="LTR50"
                     /rpt_family="LTR"
                     /rpt_type=DISPERSED
     repeat_region   268067..268247
                     /note="LTR33A"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region   269782..270098
                     /note="AluSp"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(271014..271329)
                     /note="AluJo"
                     /rpt_family="SINE/Alu"
```

```
                     /rpt_type=DISPERSED
     repeat_region   272695..272808
                     /note="AluJo"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   272809..272840
                     /note="(TAA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   272841..273006
                     /note="AluJo"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(274131..274200)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   275496..275790
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   276789..277342
                     /note="L1PA7"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(277965..278057)
                     /note="L1"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(278058..278285)
                     /note="AluSc"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(278286..278384)
                     /note="L1"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(279899..280200)
                     /note="AluSq"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   281586..281629
                     /note="(CTGGGG)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   284328..284751
                     /note="MLT1D"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   285631..285802
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   286153..286235
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
```

```
repeat_region   287050..287194
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   287532..287722
                /note="MLT1I"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   291286..291551
                /note="AluSp"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   292485..292586
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   292637..293015
                /note="L2"
                /rpt_family="LINE/L2"
                /rpt_type=DISPERSED
repeat_region   complement(293025..293207)
                /note="L1MEc"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(293208..293339)
                /note="L1PB1"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(293340..293655)
                /note="L1MEc"
                /rpt_family="LINE/L1"
                /rpt_type=DISPERSED
repeat_region   complement(293772..294060)
                /note="AluSq"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   294612..294770
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   294938..294974
                /note="(TATG)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   295016..295118
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   295484..295614
                /note="MLT1J1"
                /rpt_family="LTR/MaLR"
                /rpt_type=DISPERSED
repeat_region   296092..296159
                /note="(TG)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   complement(298406..298970)
```

```
                     /note="MLT2D"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region     complement(300518..301093)
                     /note="L1MEc"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     complement(301314..301429)
                     /note="L1MEc"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     302208..303140
                     /note="L1MC3"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     complement(303275..303473)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region     complement(303497..303797)
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     304920..305274
                     /note="AluJo"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     306916..307223
                     /note="AluSq"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region     308362..308420
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region     311366..311738
                     /note="MSTD"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region     312414..312485
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region     complement(312709..314432)
                     /note="L1MCb"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region     315703..316838
                     /note="L1ME3"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     STS               316289..316432
                     /standard_name="D21S2035"
                     /note="Accession No. G15170
                     SHGC-15782"
     repeat_region     317444..317519
                     /note="L2"
```

```
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     STS             318252..318359
                     /standard_name="A008F29"
                     /note="Accession No. R66691
                     stSG10253"
                     /db_xref="GDB:9882770"
     STS             318260..318394
                     /standard_name="sts-R67568"
     repeat_region   complement(320906..321055)
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   321167..321255
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   321719..321745
                     /note="(T)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   complement(321870..322008)
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   complement(322685..323347)
                     /note="L1"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(323404..324689)
                     /note="L1M1"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(324690..325208)
                     /note="MLT2CA"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region   complement(325209..325655)
                     /note="L1M1"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   325739..326287
                     /note="HERVL40"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region   326288..326625
                     /note="MER2"
                     /rpt_family="DNA/MER2_type"
                     /rpt_type=DISPERSED
     repeat_region   326626..326725
                     /note="HERVL40"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region   326739..326833
                     /note="HERVL40"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
```

```
repeat_region   326996..327429
                /note="LTR40b"
                /rpt_family="LTR/Retroviral"
                /rpt_type=DISPERSED
repeat_region   327971..328087
                /note="(TATATG)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   331379..331439
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   331577..331711
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   complement(332106..332264)
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   332265..332576
                /note="AluSx"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   complement(332577..332631)
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
STS             333468..333657
                /standard_name="D21S1270"
                /note="Accession No. G08079
                CHLC.GATA7G09"
repeat_region   333513..333586
                /note="(TAGA)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   333613..333682
                /note="(TG)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   complement(333703..334117)
                /note="MLT2CB"
                /rpt_family="LTR/Retroviral"
                /rpt_type=DISPERSED
repeat_region   complement(334397..334631)
                /note="MIR"
                /rpt_family="SINE/MIR"
                /rpt_type=DISPERSED
repeat_region   334655..334951
                /note="AluY"
                /rpt_family="SINE/Alu"
                /rpt_type=DISPERSED
repeat_region   338183..338215
                /note="(TTTC)n"
                /rpt_family="Simple_repeat"
                /rpt_type=TANDEM
repeat_region   complement(338216..338525)
```

```
                     /note="AluSq"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   338527..338970
                     /note="L2"
                     /rpt_family="LINE/L2"
                     /rpt_type=DISPERSED
     repeat_region   complement(339226..339529)
                     /note="AluSx"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
BASE COUNT
ORIGIN

2005 100   0.0
P_AAH64874  Human secreted protein cDNA, SEQ ID NO: 150. 045 bp,
          cDNA, PAT 11-SEP-2001
ACCESSION   P_AAH64874
KEYWORDS    GENESEQ; Human; secreted protein; gene therapy; vaccine; treatment;
          diagnosis; GENSET;  patent; patentdb (v200420, 23-SEP-2004).
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens.
REFERENCE   1  (bases 1 to 2045)
  AUTHORS   Dumas Milne Edwards,J.,  Bougueleret,L.,  Jobert,S.
  TITLE     Full length GENSET human nucleic acids encoding potentially secreted
          proteins, useful in gene therapy and vaccination against a variety
          of  diseases, and for diagnosis of those diseases.
  JOURNAL   Patent: WO200142451-A2; Filing Date: 07-DEC-2000; 2000WO-IB001938;
          Publication Date: 14-JUN-2001; Priority: 08-DEC-1999;
          99US-0169629P. 06-MAR-2000; 2000US-0187470P; Assignee: (GEST )
          GENSET; Cross Reference: WPI; 2001-367870/38. P-PSDB; AAG89271;
          Patent Format: Claim 7; Page 713-714; 921pp; English.
COMMENT     The invention relates to full length GENSET human nucleic acids
          encoding  potentially secreted proteins. The nucleic acids and the
          polypeptides  they encode may be used in the prevention, treatment
          and diagnosis of  diseases associated with inappropriate GENSET gene
          expression. For  example, they be used to treat disorders associated
          with decreased GENSET gene expression by rectifying mutations or
          deletions in a patient's  genome that affect the activity of GENSET
          or by supplementing the  patients own production of GENSET
          polypeptides. Conversely, antisense  nucleic acid molecules may be
          administered to down regulate GENSET  expression by binding with the
          cells' own genes and preventing their  expression. The sense and
          antisense nucleic acids may also be used as DNA probes in diagnostic
          assays to detect and quantitate the presence of  similar nucleic
          acid sequences in samples, and hence to determine which  patients
          may be in need of restorative therapy. The GENSET polypeptides  may
          be used as antigens in the production of antibodies and in assays to
          identify modulators (agonists and antagonists) of GENSET polypeptide
          expression and activity. The present sequence is a GENSET nucleic
          acid of the invention
FEATURES             Location/Qualifiers
BASE COUNT      620 a    352 c    398 g    675 t
ORIGIN

1930 100   0.0
HSA250711   Homo sapiens CLDN8 gene for claudin-8. 1931 bp,
```

```
          DNA, linear, PRI 11-NOV-1999
ACCESSION   AJ250711
VERSION     AJ250711.1  GI:6433859
KEYWORDS    claudin-8; CLDN8 gene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Keen,T.J. and Inglehearn,C.F.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1931)
  AUTHORS   Keen,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-1999) Keen T.J., Molecular Medicine Unit,
          University of Leeds, Clinical Sciences Building, St James's
          Hospital, Leeds, LS9 7TF, UNITED KINGDOM
FEATURES             Location/Qualifiers
     source          1..1931
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
     gene              23..700
                     /gene="CLDN8"
     CDS               23..700
                     /gene="CLDN8"
                     /codon_start=1
                     /product="claudin-8"
                     /protein_id="CAB60615.1"
                     /db_xref="GI:6433860"
                     /db_xref="GOA:P56748"
                     /db_xref="Swiss-Prot:P56748"
     polyA_signal    1913..1918
BASE COUNT
ORIGIN

1882 100   0.0
P_AAS00668  Human INTERCEPT 309 cDNA. 909 bp, cDNA, PAT 07-SEP-2001
ACCESSION   P_AAS00668
KEYWORDS    GENESEQ; Human; INTERCEPT 309; T cell; heart; liver; pancreas;
          placenta; brain; skeletal muscle; kidney; spleen; lymph node;
          peripheral blood leukocyte; bone marrow; thymus tissue; abnormal
          blood coagulation; asthma; cancer; anaphylaxis; hepatitis; multiple
          sclerosis; coronary artery disease; ss; malaria; atopic dermatitis;
          amyotrophic lateral sclerosis; meningitis; attention deficit
          disorder; Crohn's disease; gastroenteritis; goitre; hypoglycaemia;
          diabetes mellitus; endometriosis; pulmonary embolism; muscular
          dystrophy; immuno-competence; vertebrate; blood; serum; lung;
          patent; patentdb (v200420, 23-SEP-2004).
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens.
REFERENCE   1  (bases 1 to 1909)
  AUTHORS   Mackay,C.R.,  Myers,P.S.,  Kirst,S.J.,  Fraser,C.C.,  Leiby,K.R.
  TITLE     New isolated nucleic acid molecule for diagnosis, prevention, and
          therapy of human and other animal disorder, or as modulating agent
```

for regulating cellular processes.

JOURNAL Patent: WO200129088-A1; Filing Date: 23-JUN-2000; 2000WO-US017386; Publication Date: 26-APR-2001; Priority: 19-OCT-1999; 99US-00420707; Assignee: (MILL-) MILLENNIUM PHARM INC; Cross Reference: WPI; 2001-308477/32. P-PSDB; AAU00678; Patent Format: Claim 1; Fig 3A-3C; 263pp; English.

COMMENT The sequence represents a cDNA which encodes human INTERCEPT 309 polypeptide. This protein and similar others exhibit the ability to affect growth, proliferation, survival, differentiation, activity, morphology, or movement/migration of, e.g. T cells and cells of the heart, liver, pancreas, placenta, brain, lung, skeletal muscle, kidney, spleen, lymph node, peripheral blood leukocyte, bone marrow or thymus tissue. They can be used as modulating agents for regulating cellular processes, thus, the proteins and their associated nucleic acids can be used to prognosticate, prevent, diagnose, or treat disorders associated with physiological processes. These disorders include abnormal blood coagulation, asthma, anaphylaxis, hepatitis, multiple sclerosis, cancer, coronary artery disease, malaria, atopic dermatitis, amyotrophic lateral sclerosis, meningitis, attention deficit disorder, Crohn's disease, gastroenteritis, goitre, hypoglycaemia, diabetes mellitus, endometriosis, pulmonary embolism and muscular dystrophy. Antibodies to disorders such as these can be made by providing a polypeptide of the invention to an immuno-competent vertebrate and harvesting blood or serum from the vertebrate

FEATURES Location/Qualifiers

CDS 2..649
/*tag= a
/product= "Human INTERCEPT 309"

sig_peptide 2..73
/*tag= b

mat_peptide 74..646
/*tag= c
/product= "Mature human INTERCEPT 309"

BASE COUNT 586 a 318 c 368 g 637 t

ORIGIN

1832 100 0.0

P_AAH64876 Human secreted protein cDNA, SEQ ID NO: 152. 931 bp, cDNA, PAT 11-SEP-2001

ACCESSION P_AAH64876

KEYWORDS GENESEQ; Human; secreted protein; gene therapy; vaccine; treatment; diagnosis; GENSET; patent; patentdb (v200420, 23-SEP-2004).

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 1931)

AUTHORS Dumas Milne Edwards,J., Bougueleret,L., Jobert,S.

TITLE Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.

JOURNAL Patent: WO200142451-A2; Filing Date: 07-DEC-2000; 2000WO-IB001938; Publication Date: 14-JUN-2001; Priority: 08-DEC-1999; 99US-0169629P. 06-MAR-2000; 2000US-0187470P; Assignee: (GEST ) GENSET; Cross Reference: WPI; 2001-367870/38. P-PSDB; AAG89273; Patent Format: Claim 7; Page 715-716; 921pp; English.

COMMENT The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the

polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy. The GENSET polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of GENSET polypeptide expression and activity. The present sequence is a GENSET nucleic acid of the invention

FEATURES Location/Qualifiers
BASE COUNT 587 a 343 c 393 g 608 t
ORIGIN

1831 99 0.0
P_AAH15597 Human cDNA sequence SEQ ID NO:13909. 890 bp, cDNA, PAT 26-JUN-2001
ACCESSION P_AAH15597
KEYWORDS GENESEQ; Human; primer; detection; diagnosis; antisense therapy; gene therapy; patent; patentdb (v200420, 23-SEP-2004).
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1890)
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J. Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K., Otsuki,T.
TITLE Primer sets for synthesizing polynucleotides, particularly the 5602 full- length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
JOURNAL Patent: EP1074617-A2; Filing Date: 28-JUL-2000; 2000EP-00116126; Publication Date: 07-FEB-2001; Priority: 29-JUL-1999; 99JP-00248036. 27-AUG-1999; 99JP-00300253. 11-JAN-2000; 2000JP-00118776. 02-MAY-2000; 2000JP-00183767. 09-JUN-2000; 2000JP-00241899; Assignee: (HELI-) HELIX RES INST; Cross Reference: WPI; 2001-318749/34; Patent Format: Claim 8; SEQ ID NO 13909; 2537pp Sequence Listing; English.
COMMENT The present invention describes primer sets for synthesising 5602 full- length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in

antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

FEATURES Location/Qualifiers

BASE COUNT 565 a 341 c 384 g 600 t

ORIGIN

1831 99 0.0

AK022269 Homo sapiens cDNA FLJ12207 fis, clone MAMMA1000956, highly similar to Homo sapiens CLDN8 gene for claudin-8. 1890 bp, mRNA, linear, PRI 30-JAN-2004

ACCESSION AK022269

VERSION AK022269.1 GI:10433628

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.

TITLE Complete sequencing and characterization of 21,243 full-length human cDNAs

```
  JOURNAL   Nat. Genet. 36 (1), 40-45 (2004)
   PUBMED   14702039
REFERENCE   2
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
          Nishikawa,T., Nagai,K., Sato,H., Sugano,S., Shiratori,A., Sudo,H.,
          Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
          Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
          Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
          Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
          Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1890)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
          Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
          (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
          International Trade and Industry of Japan; cDNA full insert
          sequencing: Research Association for Biotechnology; cDNA library
          construction, 5'- & 3'-end one pass sequencing and clone selection:
          Helix Research Institute (supported by Japan Key Technology Center
          etc.) and Department of Virology, Institute of Medical Science,
          University of Tokyo.
FEATURES             Location/Qualifiers
     source          1..1890
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MAMMA1000956"
                     /tissue_type="Mammary gland"
                     /clone_lib="MAMMA1"
                     /note="cloning vector: pME18SFL3"
     misc_feature    795^796
                     /note="5 bases segment is present in AJ250711 , AP001707.1
                     and AP000884.1."
BASE COUNT
ORIGIN

1692 100   0.0
P_AAH64888  Human secreted protein cDNA, SEQ ID NO: 164. 793 bp,
          cDNA, PAT 11-SEP-2001
ACCESSION   P_AAH64888
KEYWORDS    GENESEQ; Human; secreted protein; gene therapy; vaccine; treatment;
          diagnosis; GENSET;  patent; patentdb (v200420, 23-SEP-2004).
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens.
REFERENCE   1  (bases 1 to 1793)
  AUTHORS   Dumas Milne Edwards,J.,  Bougueleret,L.,  Jobert,S.
  TITLE     Full length GENSET human nucleic acids encoding potentially secreted
          proteins, useful in gene therapy and vaccination against a variety
          of  diseases, and for diagnosis of those diseases.
  JOURNAL   Patent: WO200142451-A2; Filing Date: 07-DEC-2000; 2000WO-IB001938;
          Publication Date: 14-JUN-2001; Priority: 08-DEC-1999;
          99US-0169629P. 06-MAR-2000; 2000US-0187470P; Assignee: (GEST )
          GENSET; Cross Reference: WPI; 2001-367870/38. P-PSDB; AAG89285;
```

Patent Format: Claim 7; Page 726-727; 921pp; English.

COMMENT The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy. The GENSET polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of GENSET polypeptide expression and activity. The present sequence is a GENSET nucleic acid of the invention

FEATURES Location/Qualifiers

BASE COUNT 544 a 318 c 369 g 556 t

ORIGIN

1689 99 0.0

P_AAH64890 Human secreted protein cDNA, SEQ ID NO: 166. 748 bp, cDNA, PAT 11-SEP-2001

ACCESSION P_AAH64890

KEYWORDS GENESEQ; Human; secreted protein; gene therapy; vaccine; treatment; diagnosis; GENSET; patent; patentdb (v200420, 23-SEP-2004).

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 1748)

AUTHORS Dumas Milne Edwards,J., Bougueleret,L., Jobert,S.

TITLE Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.

JOURNAL Patent: WO200142451-A2; Filing Date: 07-DEC-2000; 2000WO-IB001938; Publication Date: 14-JUN-2001; Priority: 08-DEC-1999; 99US-0169629P. 06-MAR-2000; 2000US-0187470P; Assignee: (GEST ) GENSET; Cross Reference: WPI; 2001-367870/38. P-PSDB; AAG89287; Patent Format: Claim 7; Page 729; 921pp; English.

COMMENT The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy. The GENSET polypeptides may

be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of GENSET polypeptide expression and activity. The present sequence is a GENSET nucleic acid of the invention

FEATURES Location/Qualifiers

BASE COUNT 530 a 311 c 360 g 545 t

ORIGIN

# Genome-wide Study of Gene Copy Numbers, Transcripts, and Protein Levels in Pairs of Non-invasive and Invasive Human Transitional Cell Carcinomas*

Torben F. Ørntoft‡§, Thomas Thykjaer¶, Frederic M. Waldman‖, Hans Wolf**, and Julio E. Celis‡‡

**Gain and loss of chromosomal material is characteristic of bladder cancer, as well as malignant transformation in general. The consequences of these changes at both the transcription and translation levels is at present unknown partly because of technical limitations. Here we have attempted to address this question in pairs of non-invasive and invasive human bladder tumors using a combination of technology that included comparative genomic hybridization, high density oligonucleotide array-based monitoring of transcript levels (5600 genes), and high resolution two-dimensional gel electrophoresis. The results showed that there is a gene dosage effect that in some cases superimposes on other regulatory mechanisms. This effect depended ($p < 0.015$) on the magnitude of the comparative genomic hybridization change. In general (18 of 23 cases), chromosomal areas with more than 2-fold gain of DNA showed a corresponding increase in mRNA transcripts. Areas with loss of DNA, on the other hand, showed either reduced or unaltered transcript levels. Because most proteins resolved by two-dimensional gels are unknown it was only possible to compare mRNA and protein alterations in relatively few cases of well focused abundant proteins. With few exceptions we found a good correlation ($p < 0.005$) between transcript alterations and protein levels. The implications, as well as limitations, of the approach are discussed.** ***Molecular & Cellular Proteomics 1:37–45, 2002.***

Aneuploidy is a common feature of most human cancers (1), but little is known about the genome-wide effect of this phenomenon at both the transcription and translation levels. High throughput array studies of the breast cancer cell line BT474 has suggested that there is a correlation between DNA copy numbers and gene expression in highly amplified areas (2), and studies of individual genes in solid tumors have revealed a good correlation between gene dose and mRNA or protein levels in the case of c-erb-B2, *cyclin d1*, *ems1*, and N-myc (3–5). However, a high cyclin D1 protein expression has been observed without simultaneous amplification (4), and a low level of c-myc copy number increase was observed without concomitant c-myc protein overexpression (6).

In human bladder tumors, karyotyping, fluorescent *in situ* hybridization, and comparative genomic hybridization (CGH)[1] have revealed chromosomal aberrations that seem to be characteristic of certain stages of disease progression. In the case of non-invasive pTa transitional cell carcinomas (TCCs), this includes loss of chromosome 9 or parts of it, as well as loss of Y in males. In minimally invasive pT1 TCCs, the following alterations have been reported: 2q−, 11p−, 1q+, 11q13+, 17q+, and 20q+ (7–12). It has been suggested that these regions harbor tumor suppressor genes and oncogenes; however, the large chromosomal areas involved often contain many genes, making meaningful predictions of the functional consequences of losses and gains very difficult.

In this investigation we have combined genome-wide technology for detecting genomic gains and losses (CGH) with gene expression profiling techniques (microarrays and proteomics) to determine the effect of gene copy number on transcript and protein levels in pairs of non-invasive and invasive human bladder TCCs.

## EXPERIMENTAL PROCEDURES

*Material*—Bladder tumor biopsies were sampled after informed consent was obtained and after removal of tissue for routine pathology examination. By light microscopy tumors 335 and 532 were staged by an experienced pathologist as pTa (superficial papillary),

From the ‡Department of Clinical Biochemistry, Molecular Diagnostic Laboratory and **Department of Urology, Aarhus University Hospital, Skejby, DK-8200 Aarhus N, Denmark, ¶AROS Applied Biotechnology ApS, Gustav Wiedsvej 10, DK-8000 Aarhus C, Denmark, ‖UCSF Cancer Center and Department of Laboratory Medicine, University of California, San Francisco, CA 94143-0808, and ‡‡Institute of Medical Biochemistry and Danish Centre for Human Genome Research, Ole Worms Allé 170, Aarhus University, DK-8000 Aarhus C, Denmark

Received, September 26, 2001, and in revised form, November 7, 2001

Published, MCP Papers in Press, November 13, 2001, DOI 10.1074/mcp.M100019-MCP200

[1] The abbreviations used are: CGH, comparative genomic hybridization; TCC, transitional cell carcinoma; LOH, loss of heterozygosity; PA-FABP, psoriasis-associated fatty acid-binding protein; 2D, two-dimensional.

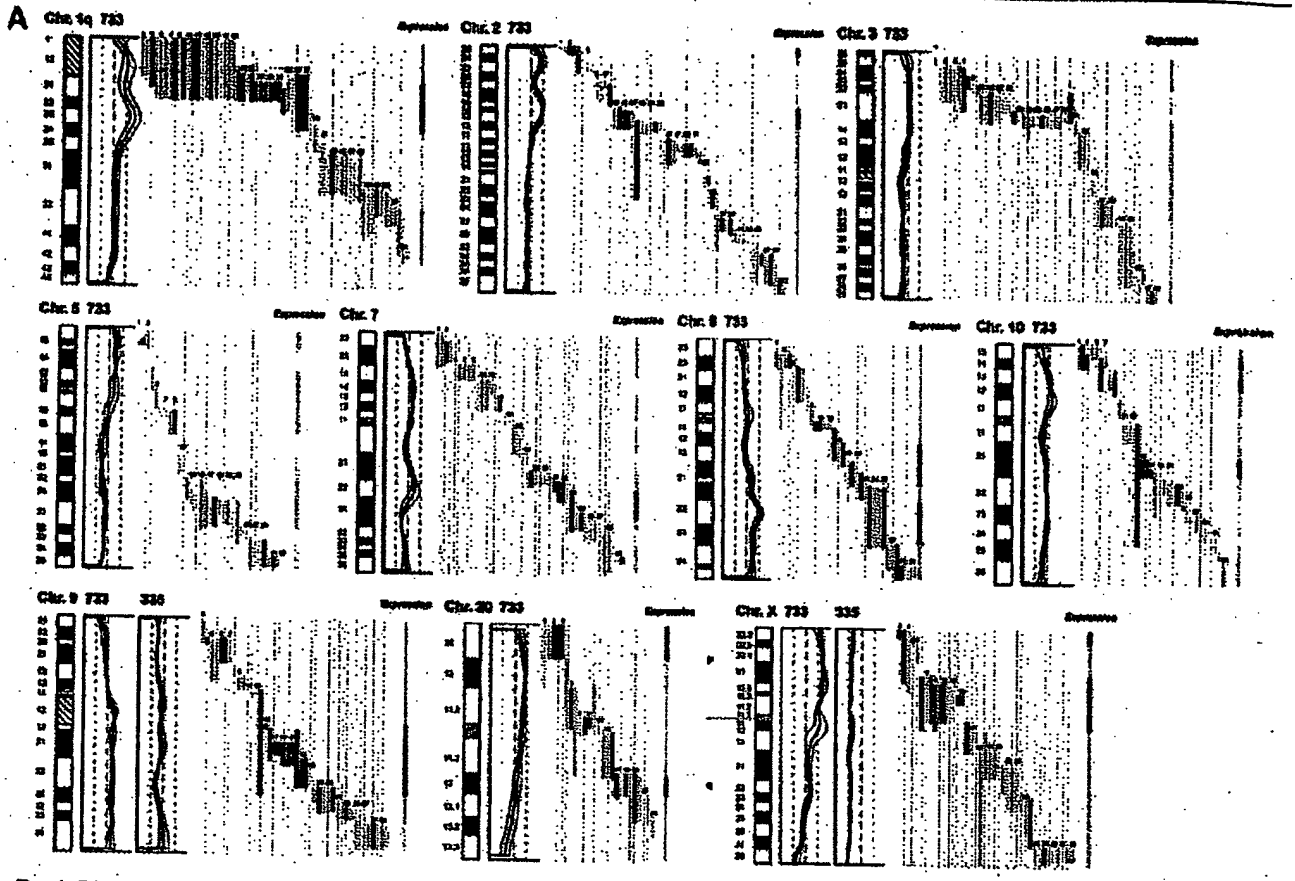



FIG. 1. **DNA copy number and mRNA expression level.** Shown from *left* to *right* are chromosome (*Chr.*), CGH profiles, gene location and expression level of specific genes, and overall expression level along the chromosome. *A*, expression of mRNA in invasive tumor 733 as compared with the non-invasive counterpart tumor 335. *B*, expression of mRNA in invasive tumor 827 compared with the non-invasive counterpart tumor 532. The average fluorescent signal ratio between tumor DNA and normal DNA is shown along the length of the chromosome (*left*). The *bold curve* in the ratio profile represents a mean of four chromosomes and is surrounded by *thin curves* indicating one standard deviation. The *central vertical line* (*broken*) indicates a ratio value of 1 (no change), and the *vertical lines* next to it (*dotted*) indicate a ratio of 0.5 (*left*) and 2.0 (*right*). In chromosomes where the non-invasive tumor 335 used for comparison showed alterations in DNA content, the ratio profile of that chromosome is shown to the *right* of the invasive tumor profile. The *colored bars* represents one gene each, identified by the running *numbers above* the *bars* (the name of the gene can be seen at www.MDL.DK/sdata.html). The *bars* indicate the purported location of the gene, and the *colors* indicate the expression level of the gene in the invasive tumor compared with the non-invasive counterpart; >2-fold increase (*black*), >2-fold decrease (*blue*), no significant change (*orange*). The *bar* to the *far right*, entitled *Expression* shows the resulting change in expression along the chromosome; the *colors* indicate that at least half of the genes were up-regulated (*black*), at least half of the genes down-regulated (*blue*), or more than half of the genes are unchanged (*orange*). If a gene was absent in one of the samples and present in another, it was regarded as more than a 2-fold change. A 2-fold level was chosen as this corresponded to one standard deviation in a double determination of ~1800 genes. Centromeres and heterochromatic regions were excluded from data analysis.

grade I and II, respectively, tumors 733 and 827 were staged as pT1 (invasive into submucosa), 733 was staged as solid, and 827 was staged as papillary, both grade III.

*mRNA Preparation*—Tissue biopsies, obtained fresh from surgery, were embedded immediately in a sodium-guanidinium thiocyanate solution and stored at −80 °C. Total RNA was isolated using the RNAzol B RNA isolation method (WAK-Chemie Medical GMBH). poly(A)$^+$ RNA was isolated by an oligo(dT) selection step (Oligotex mRNA kit; Qiagen).

*cRNA Preparation*—1 μg of mRNA was used as starting material. The first and second strand cDNA synthesis was performed using the SuperScript® choice system (Invitrogen) according to the manufacturer's instructions but using an oligo(dT) primer containing a T7 RNA polymerase binding site. Labeled cRNA was prepared using the MEGAscrip® *in vitro* transcription kit (Ambion). Biotin-labeled CTP and UTP (Enzo) was used, together with unlabeled NTPs in the reaction. Following the *in vitro* transcription reaction, the unincorporated nucleotides were removed using RNeasy columns (Qiagen).

*Array Hybridization and Scanning*—Array hybridization and scanning was modified from a previous method (13). 10 μg of cRNA was fragmented at 94 °C for 35 min in buffer containing 40 mM Tris acetate, pH 8.1, 100 mM KOAc, 30 mM MgOAc. Prior to hybridization, the fragmented cRNA in a 6× SSPE-T hybridization buffer (1 M NaCl, 10 mM Tris, pH 7.6, 0.005% Triton), was heated to 95 °C for 5 min, subsequently cooled to 40 °C, and loaded onto the Affymetrix probe array cartridge. The probe array was then incubated for 16 h at 40 °C at constant rotation (60 rpm). The probe array was exposed to 10 washes in 6× SSPE-T at 25 °C followed by 4 washes in 0.5× SSPE-T at 50 °C. The biotinylated cRNA was stained with a streptavidin-phycoerythrin conjugate, 10 μg/ml (Molecular Probes) in 6× SSPE-T

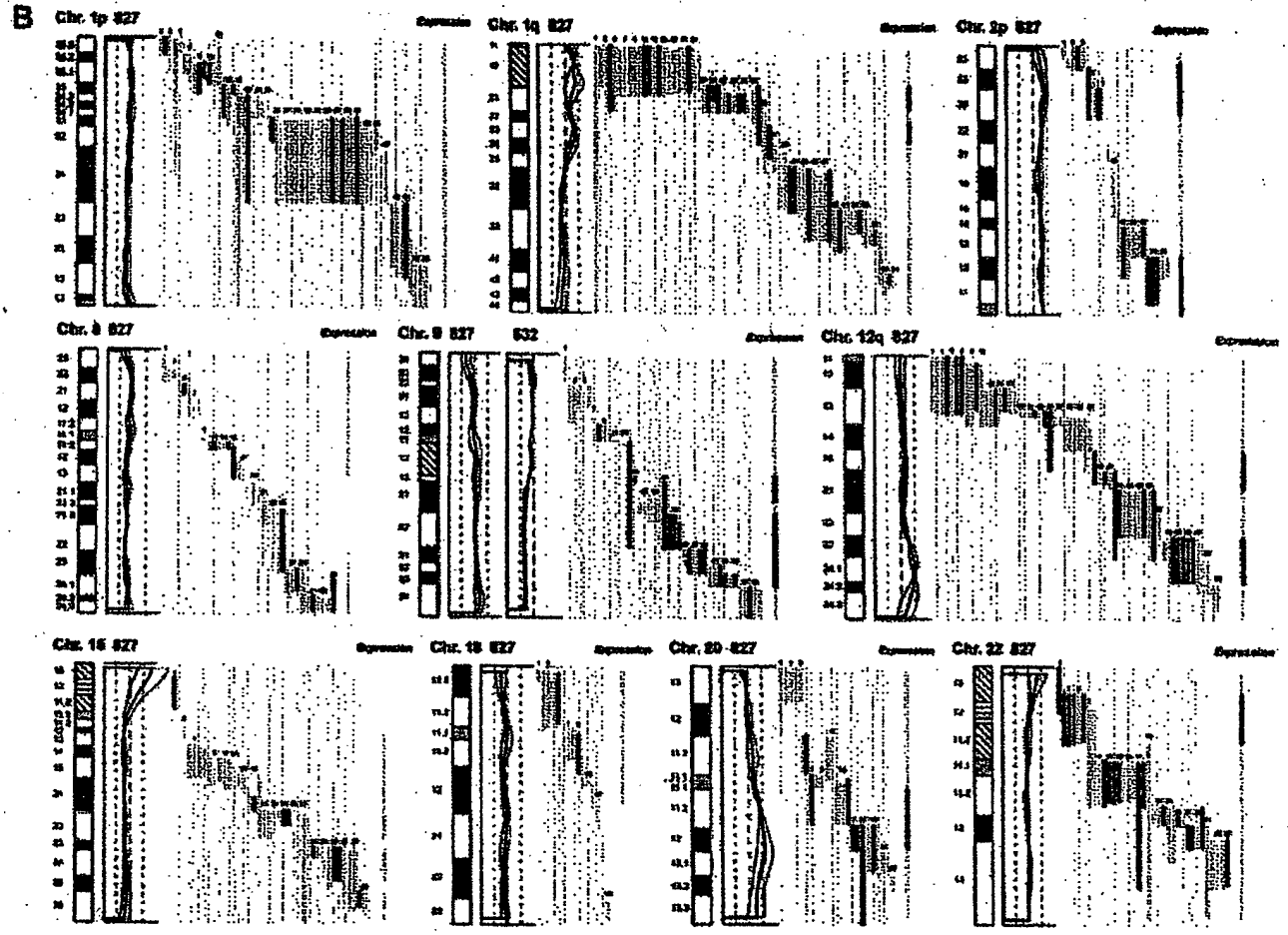



FIG. 1—*continued*

for 30 min at 25 °C followed by 10 washes in 6× SSPE-T at 25 °C. The probe arrays were scanned at 560 nm using a confocal laser scanning microscope (made for Affymetrix by Hewlett-Packard). The readings from the quantitative scanning were analyzed by Affymetrix gene expression analysis software.

*Microsatellite Analysis*—Microsatellite Analysis was performed as described previously (14). Microsatellites were selected by use of www.nobi.nlm.nih.gov/genemap98, and primer sequences were obtained from the genome data base at www.gdb.org. DNA was extracted from tumor and blood and amplified by PCR in a volume of 20 μl for 35 cycles. The amplicons were denatured and electrophoresed for 3 h in an ABI Prism 377. Data were collected in the Gene Scan program for fragment analysis. Loss of heterozygosity was defined as less than 33% of one allele detected in tumor amplicons compared with blood.

*Proteomic Analysis*—TCCs were minced into small pieces and homogenized in a small glass homogenizer in 0.5 ml of lysis solution. Samples were stored at −20 °C until use. The procedure for 2D gel electrophoresis has been described in detail elsewhere (15, 16). Gels were stained with silver nitrate and/or Coomassie Brilliant Blue. Proteins were identified by a combination of procedures that included microsequencing, mass spectrometry, two-dimensional gel Western immunoblotting, and comparison with the master two-dimensional gel image of human keratinocyte proteins; see biobase.dk/cgi-bin/cells.

*CGH*—Hybridization of differentially labeled tumor and normal DNA to normal metaphase chromosomes was performed as described previously (10). Fluorescein-labeled tumor DNA (200 ng), Texas Red-labeled reference DNA (200 ng), and human Cot-1 DNA (20 μg) were denatured at 37 °C for 5 min and applied to denatured normal metaphase slides. Hybridization was at 37 °C for 2 days. After washing, the slides were counterstained with 0.15 μg/ml 4,6-diamidino-2-phenylindole in an anti-fade solution. A second hybridization was performed for all tumor samples using fluorescein-labeled reference DNA and Texas Red-labeled tumor DNA (inverse labeling) to confirm the aberrations detected during the initial hybridization. Each CGH experiment also included a normal control hybridization using fluorescein- and Texas Red-labeled normal DNA. Digital image analysis was used to identify chromosomal regions with abnormal fluorescence ratios, indicating regions of DNA gains and losses. The average green:red fluorescence intensity ratio profiles were calculated using four images of each chromosome (eight chromosomes total) with normalization of the green:red fluorescence intensity ratio for the entire metaphase and background correction. Chromosome identification was performed based on 4,6-diamidino-2-phenylindole banding patterns. Only images showing uniform high intensity fluorescence with minimal background staining were analyzed. All centromeres, p arms of acrocentric chromosomes, and heterochromatic regions were excluded from the analysis.

## RESULTS

*Comparative Genomic Hybridization*—The CGH analysis identified a number of chromosomal gains and losses in the

TABLE I

*Correlation between alterations detected by CGH and by expression monitoring*

Top, CGH used as independent variable (if CGH alteration – what expression ratio was found); bottom, altered expression used as independent variable (if expression alteration – what CGH deviation was found).

| CGH alterations | Tumor 733 vs. 335 Expression change clusters | Concordance | CGH alterations | Tumor 827 vs. 532 Expression change clusters | Concordance |
|---|---|---|---|---|---|
| 13 Gain | 10 Up-regulation<br>0 Down-regulation<br>3 No change | 77% | 10 Gain | 8 Up-regulation<br>0 Down-regulation<br>2 No change | 80% |
| 10 Loss | 1 Up-regulation<br>5 Down-regulation<br>4 No change | 50% | 12 Loss | 3 Up-regulation<br>2 Down regulation<br>7 No change | 17% |

| Expression change clusters | Tumor 733 vs. 335 CGH alterations | Concordance | Expression change clusters | Tumor 827 vs. 532 CGH alterations | Concordance |
|---|---|---|---|---|---|
| 16 Up-regulation | 11 Gain<br>2 Loss<br>3 No change | 69% | 17 Up-regulation | 10 Gain<br>5 Loss<br>2 No change | 59% |
| 21 Down-regulation | 1 Gain<br>8 Loss<br>12 No change | 38% | 9 Down-regulation | 0 Gain<br>3 Loss<br>6 No change | 33% |
| 15 No change | 3 Gain<br>3 Loss<br>9 No change | 60% | 21 No change | 1 Gain<br>3 Loss<br>17 No change | 81% |

two invasive tumors (stage pT1, TCCs 733 and 827), whereas the two non-invasive papillomas (stage pTa, TCCs 335 and 532) showed only 9p–, 9q22-q33–, and X–, and 7+, 9q–, and Y–, respectively. Both invasive tumors showed changes (1q22-24+, 2q14.1-qter–, 3q12-q13.3–, 6q12-q22–, 9q34+, 11q12-q13+, 17+, and 20q11.2-q12+) that are typical for their disease stage, as well as additional alterations, some of which are shown in Fig. 1. Areas with gains and losses deviated from the normal copy number to some extent, and the average numerical deviation from normal was 0.4-fold in the case of TCC 733 and 0.3-fold for TCC 827. The largest changes, amounting to at least a doubling of chromosomal content, were observed at 1q23 in TCC 733 (Fig. 1*A*) and 20q12 in TCC 827 (Fig. 1*B*).

*mRNA Expression in Relation to DNA Copy Number*—The mRNA levels from the two invasive tumors (TCCs 827 and 733) were compared with the two non-invasive counterparts (TCCs 532 and 335). This was done in two separate experiments in which we compared TCCs 733 to 335 and 827 to 532, respectively, using two different scaling settings for the arrays to rule out scaling as a confounding parameter. Approximately 1,800 genes that yielded a signal on the arrays were searched in the Unigene and Genemap data bases for chromosomal location, and those with a known location (1096) were plotted as bars covering their purported locus. In that way it was possible to construct a graphic presentation of DNA copy number and relative mRNA levels along the individual chromosomes (Fig. 1).

For each mRNA a ratio was calculated between the level in the invasive *versus* the non-invasive counterpart. Bars, which represent chromosomal location of a gene, were color-coded according to the expression ratio, and only differences larger than 2-fold were regarded as informative (Fig. 1). The density of genes along the chromosomes varied, and areas containing only one gene were excluded from the calculations. The resolution of the CGH method is very low, and some of the outlier data may be because of the fact that the boundaries of the chromosomal aberrations are not known at high resolution.

Two sets of calculations were made from the data. For the first set we used CGH alterations as the independent variable and estimated the frequency of expression alterations in these chromosomal areas. In general, areas with a strong gain of chromosomal material contained a cluster of genes having increased mRNA expression. For example, both chromosomes 1q21-q25, 2p and 9q, showed a relative gain of more than 100% in DNA copy number that was accompanied by increased mRNA expression levels in the two tumor pairs (Fig. 1). In most cases, chromosomal gains detected by CGH were accompanied by an increased level of transcripts in both TCCs 733 (77%) and 827 (80%) (Table I, *top*). Chromosomal losses, on the other hand, were not accompanied by decreased expression in several cases, and were often registered as having unaltered RNA levels (Table I, *top*). The inability to detect RNA expression changes in these cases was not because of fewer genes mapping to the lost regions (data not shown).

In the second set of calculations we selected expression alterations above 2-fold as the independent variable and estimated the frequency of CGH alterations in these areas. As above, we found that increased transcript expression correlated with gain of chromosomal material (TCC 733, 69% and TCC 827, 59%), whereas reduced expression was often detected in areas with unaltered CGH ratios (Table I, *bottom*). Furthermore, as a control we looked at areas with no alter-

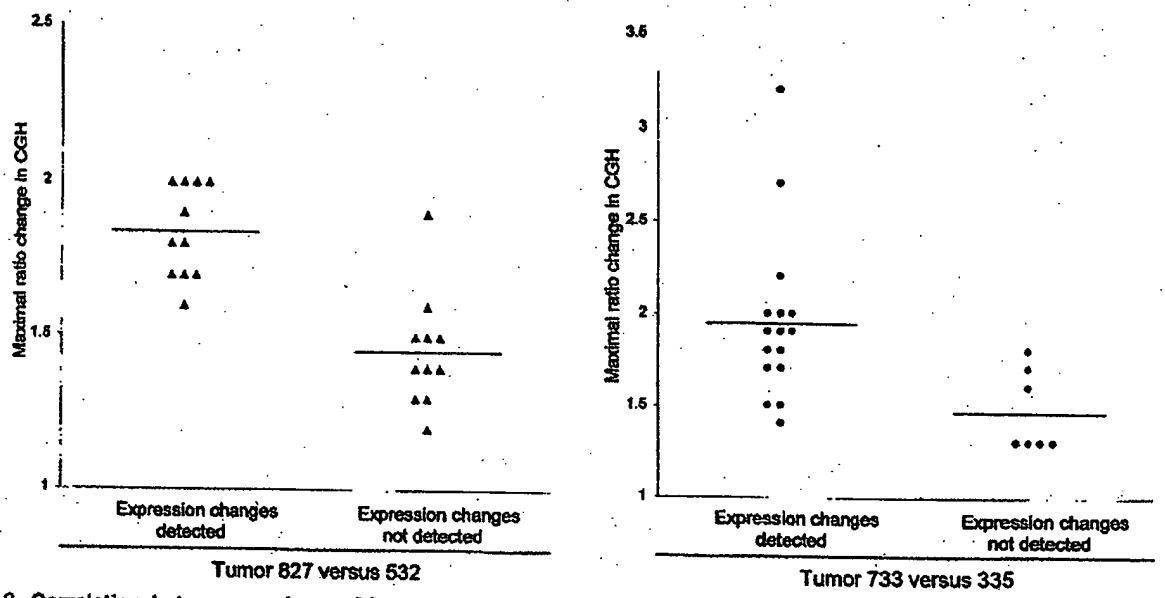



FIG. 2. **Correlation between maximum CGH aberration and the ability to detect expression change by oligonucleotide array monitoring.** The aberration is shown as a numerical -fold change in ratio between invasive tumors 827 (▲) and 733 (♦) and their non-invasive counterparts 532 and 335. The expression change was taken from the *Expression* line to the *right* in Fig. 1, which depicts the resulting expression change for a given chromosomal region. At least half of the mRNAs from a given region have to be either up- or down-regulated to be scored as an expression change. All chromosomal arms in which the CGH ratio plus or minus one standard deviation was outside the ratio value of one were included.

ation in expression. No alteration was detected by CGH in most of these areas (TCC 733, 60% and TCC 827, 81%; see Table I, *bottom*). Because the ability to observe reduced or increased mRNA expression clustering to a certain chromosomal area clearly reflected the extent of copy number changes, we plotted the maximum CGH aberrations in the regions showing CGH changes against the ability to detect a change in mRNA expression as monitored by the oligonucleotide arrays (Fig. 2). For both tumors TCC 733 ($p < 0.015$) and TCC 827 ($p < 0.00003$) a highly significant correlation was observed between the level of CGH ratio change (reflecting the DNA copy number) and alterations detected by the array based technology (Fig. 2). Similar data were obtained when areas with altered expression were used as independent variables. These areas correlated best with CGH when the CGH ratio deviated 1.6- to 2.0-fold (Table I, *bottom*) but mostly did not at lower CGH deviations. These data probably reflect that loss of an allele may only lead to a 50% reduction in expression level, which is at the cut-off point for detection of expression alterations. Gain of chromosomal material can occur to a much larger extent.

*Microsatellite-based Detection of Minor Areas of Losses*—In TCC 733, several chromosomal areas exhibiting DNA amplification were preceded or followed by areas with a normal CGH but reduced mRNA expression (see Fig. 1, TCC 733 chromosome 1q32, 2p21, and 7q21 and q32, 9q34, and 10q22). To determine whether these results were because of undetected loss of chromosomal material in these regions or because of other non-structural mechanisms regulating transcription, we examined two microsatellites positioned at chromosome 1q25–32 and two at chromosome 2p22. Loss of heterozygosity (LOH) was found at both 1q25 and at 2p22 indicating that minor deleted areas were not detected with the resolution of CGH (Fig. 3). Additionally, chromosome 2p in TCC 733 showed a CGH pattern of gain/no change/gain of DNA that correlated with transcript increase/decrease/increase. Thus, for the areas showing increased expression there was a correlation with the DNA copy number alterations (Fig. 1*A*). As indicated above, the mRNA decrease observed in the middle of the chromosomal gain was because of LOH, implying that one of the mechanisms for mRNA down-regulation may be regions that have undergone smaller losses of chromosomal material. However, this cannot be detected with the resolution of the CGH method.

In both TCC 733 and TCC 827, the telomeric end of chromosome 11p showed a normal ratio in the CGH analysis; however, clusters of five and three genes, respectively, lost their expression. Two microsatellites (D11S1760, D11S922) positioned close to MUC2, IGF2, and cathepsin D indicated LOH as the most likely mechanism behind the loss of expression (data not shown).

A reduced expression of mRNA observed in TCC 733 at chromosomes 3q24, 11p11, 12p12.2, 12q21.1, and 16q24 and in TCC 827 at chromosome 11p15.5, 12p11, 15q11.2, and 18q12 was also examined for chromosomal losses using microsatellites positioned as close as possible to the gene loci

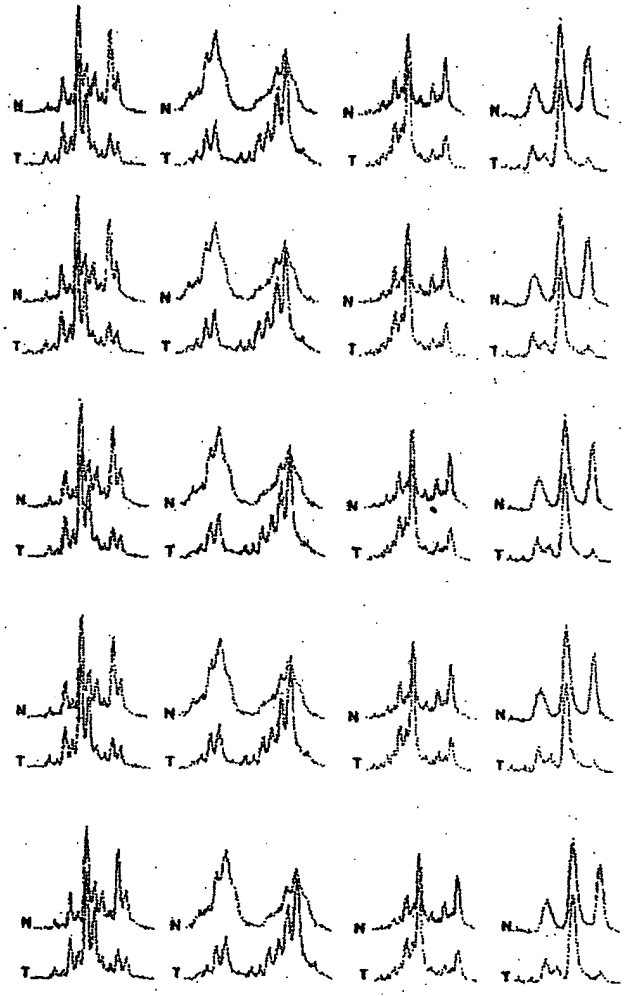



FIG. 3. **Microsatellite analysis of loss of heterozygosity.** Tumor 733 showing loss of heterozygosity at chromosome 1q25, detected (*a*) by D1S215 close to Hu class I histocompatibility antigen (gene number 38 in Fig. 1), (*b*) by D1S2735 close to cathepsin E (gene number 41 in Fig. 1), and (*c*) at chromosome 2p23 by D2S2251 close to general β-spectrin (gene number 11 on Fig. 1) and of (*d*) tumor 827 showing loss of heterozygosity at chromosome 18q12 by S18S1118 close to mitochondrial 3-oxoacyl-coenzyme A thiolase (gene number 12 in Fig. 1). The *upper curves* show the electropherogram obtained from normal DNA from leukocytes (*N*), and the *lower curves* show the electropherogram from tumor DNA (*T*). In all cases one allele is partially lost in the tumor amplicon.

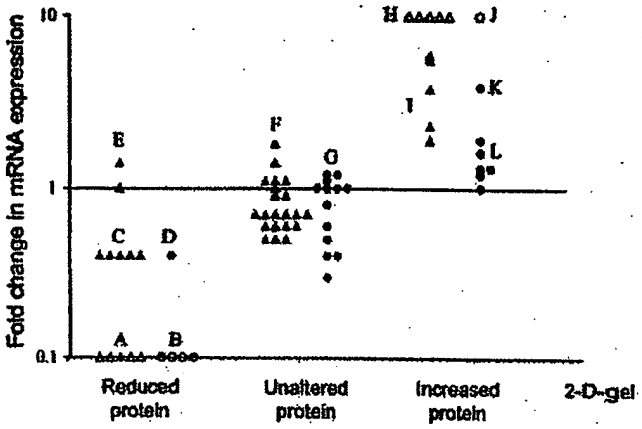



FIG. 4. **Correlation between protein levels as judged by 2D-PAGE and transcript ratio.** For comparison proteins were divided in three groups, unaltered in level or up- or down-regulated (*horizontal axis*). The mRNA ratio as determined by oligonucleotide arrays was plotted for each gene (*vertical axis*). ▲, mRNAs that were scored as present in both tumors used for the ratio calculation; △, mRNAs that were scored as absent in the invasive tumors (along *horizontal axis*) or as absent in non-invasive reference (*top* of figure). Two different scalings were used to exclude scaling as a confounder, TCCs 827 and 532 (▲△) were scaled with background suppression, and TCCs 733 and 335 (●○) were scaled without suppression. Both comparisons showed highly significant ($p < 0.005$) differences in mRNA ratios between the groups. Proteins shown were as follows: Group *A* (from *left*), phosphoglucomutase 1, glutathione transferase class μ number 4, fatty acid-binding protein homologue, cytokeratin 15, and cytokeratin 13; *B* (from *left*), fatty acid-binding protein homologue, 28-kDa heat shock protein, cytokeratin 13, and calcyclin; *C* (from *left*), α-enolase, hnRNP B1, 28-kDa heat shock protein, 14-3-3-ε, and pre-mRNA splicing factor; *D*, mesothelial keratin K7 (type II); *E* (from *top*), glutathione *S*-transferase-π and mesothelial keratin K7 (type II); *F* (from *top* and *left*), adenylyl cyclase-associated protein, E-cadherin, keratin 19, calgizzarin, phosphoglycerate mutase, annexin IV, cytoskeletal γ-actin, hnRNP A1, integral membrane protein calnexin (IP90), hnRNP H, brain-type clathrin light chain-a, hnRNP F, 70-kDa heat shock protein, heterogeneous nuclear ribonucleoprotein A/B, translationally controlled tumor protein, liver glyceraldehyde-3-phosphate dehydrogenase, keratin 8, aldehyde reductase, and Na,K-ATPase β-1 subunit; *G*, (from *top* and *left*), TCP20, calgizzarin, 70-kDa heat shock protein, calnexin, hnRNP H, cytokeratin 15, ATP synthase, keratin 19, triosephosphate isomerase, hnRNP F, liver glyceraldehyde-3-phosphatase dehydrogenase, glutathione *S*-transferase-π, and keratin 8; *H* (from *left*), plasma gelsolin, autoantigen calreticulin, thioredoxin, and NAD+-dependent 15 hydroxyprostaglandin dehydrogenase; *I* (from *top*), prolyl 4-hydroxylase β-subunit, cytokeratin 20, cytokeratin 17, prohibition, and fructose 1,6-biphosphatase; *J* annexin II; *K*, annexin IV; *L* (from *top* and *left*), 90-kDa heat shock protein, prolyl 4-hydroxylase β-subunit, α-enolase, GRP 78, cyclophilin, and cofilin.

showing reduced mRNA transcripts. Only the microsatellite positioned at 18q12 showed LOH (Fig. 3), suggesting that transcriptional down-regulation of genes in the other regions may be controlled by other mechanisms.

*Relation between Changes in mRNA and Protein Levels*—2D-PAGE analysis, in combination with Coomassie Brilliant Blue and/or silver staining, was carried out on all four tumors using fresh biopsy material. 40 well resolved abundant known proteins migrating in areas away from the edges of the pH gradient, and having a known chromosomal location, were selected for analysis in the TCC pair 827/532. Proteins were identified by a combination of methods (see "Experimental Procedures"). In general there was a highly significant correlation ($p < 0.005$) between mRNA and protein alterations (Fig. 4). Only one gene showed disagreement between transcript alteration and protein alteration. Except for a group of cyto-

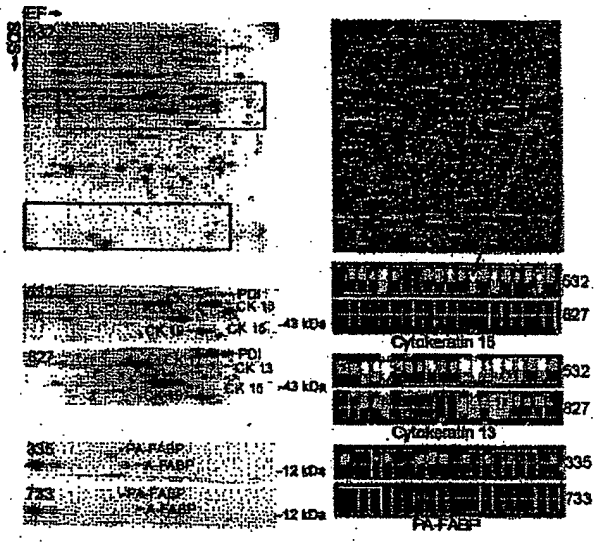



FIG. 5. **Comparison of protein and transcript levels in invasive and non-invasive TCCs.** The *upper part* of the figure shows a 2D gel (*left*) and the oligonucleotide array (*right*) of TCC 532. The *red rectangles* on the *upper gel* highlight the areas that are compared below. Identical areas of 2D gels of TCCs 532 and 827 are shown below. Clearly, cytokeratins 13 and 15 are strongly down-regulated in TCC 827 (*red annotation*). The tile on the array containing probes for cytokeratin 15 is enlarged *below* the array (*red arrow*) from TCC 532 and is compared with TCC 827. The *upper row* of *squares* in each tile corresponds to perfect match probes; the *lower row* corresponds to mismatch probes containing a mutation (used for correction for unspecific binding). Absence of signal is depicted as *black*, and the higher the signal the lighter the color. A high transcript level was detected in TCC 532 (6151 units) whereas a much lower level was detected in TCC 827 (absence of signals). For cytokeratin 13, a high transcript level was also present in TCC 532 (15659 units), and a much lower level was present in TCC 827 (623 units). The 2D gels at the *bottom* of the figure (*left*) show levels of PA-FABP and adipocyte-FABP in TCCs 335 and 733 (invasive), respectively. Both proteins are down-regulated in the invasive tumor. To the *right* we show the array tiles for the PA-FABP transcript. A medium transcript level was detected in the case of TCC 335 (1277 units) whereas very low levels were detected in TCC 733 (166 units). *IEF*, isoelectric focusing.

keratins encoded by genes on chromosome 17 (Fig. 5) the analyzed proteins did not belong to a particular family. 26 well focused proteins whose genes had a know chromosomal location were detected in TCCs 733 and 335, and of these 19 correlated ($p < 0.005$) with the mRNA changes detected using the arrays (Fig. 4). For example, PA-FABP was highly expressed in the non-invasive TCC 335 but lost in the invasive counterpart (TCC 733; see Fig. 5). The smaller number of proteins detected in both 733 and 335 was because of the smaller size of the biopsies that were available.

11 chromosomal regions where CGH showed aberrations that corresponded to the changes in transcript levels also showed corresponding changes in the protein level (Table II). These regions included genes that encode proteins that are found to be frequently altered in bladder cancer, namely cytokeratins 17 and 20, annexins II and IV, and the fatty acid-binding proteins PA-FABP and FBP1. Four of these proteins were encoded by genes in chromosome 17q, a frequently amplified chromosomal area in invasive bladder cancers.

## DISCUSSION

Most human cancers have abnormal DNA content, having lost some chromosomal parts and gained others. The present study provides some evidence as to the effect of these gains and losses on gene expression in two pairs of non-invasive and invasive TCCs using high throughput expression arrays and proteomics, in combination with CGH. In general, the results showed that there is a clear individual regulation of the mRNA expression of single genes, which in some cases was superimposed by a DNA copy number effect. In most cases, genes located in chromosomal areas with gains often exhibited increased mRNA expression, whereas areas showing losses showed either no change or a reduced mRNA expression. The latter might be because of the fact that losses most often are restricted to loss of one allele, and the cut-off point for detection of expression alterations was a 2-fold change, thus being at the border of detection. In several cases, how-

TABLE II
*Proteins whose expression level correlates with both mRNA and gene dose changes*

| Protein | Chromosomal location | Tumor TCC | CGH alteration | Transcript alteration[a] | Protein alteration |
|---|---|---|---|---|---|
| Annexin II | 1q21 | 733 | Gain | Abs to Pres[a] | Increase |
| Annexin IV | 2p13 | 733 | Gain | 3.9-Fold up | Increase |
| Cytokeratin 17 | 17q12-q21 | 827 | Gain | 3.8-Fold up | Increase |
| Cytokeratin 20 | 17q21.1 | 827 | Gain | 5.6-Fold up | Increase |
| (PA-)FABP | 8q21.2 | 827 | Loss | 10-Fold down | Decrease |
| FBP1 | 9q22 | 827 | Gain | 2.3-Fold up | Increase |
| Plasma gelsolin | 9q31 | 827 | Gain | Abs to Pres | Increase |
| Heat shock protein 28 | 15q12-q13 | 827 | Loss | 2.5-Fold up | Decrease |
| Prohibitin | 17q21 | 827/733 | Gain | 3.7-/2.5-Fold up[b] | Increase |
| Prolyl-4-hydroxyl | 17q25 | 827/733 | Gain | 5.7-/1.6-Fold up | Increase |
| hnRNPB1 | 7p15 | 827 | Loss | 2.5-Fold down | Decrease |

[a] Abs, absent; Pres, present.
[b] In cases where the corresponding alterations were found in both TCCs 827 and 733 these are shown as 827/733.

ever, an increase or decrease in DNA copy number was associated with *de novo* occurrence or complete loss of transcript, respectively. Some of these transcripts could not be detected in the non-invasive tumor but were present at relatively high levels in areas with DNA amplifications in the invasive tumors (e.g. in TCC 733 transcript from cellular ligand of annexin II gene (chromosome 1q21) from absent to 2670 arbitrary units; in TCC 827 transcript from small proline-rich protein 1 gene (chromosome 1q12-q21.1) from absent to 1326 arbitrary units). It may be anticipated from these data that significant clustering of genes with an increased expression to a certain chromosomal area indicates an increased likelihood of gain of chromosomal material in this area.

Considering the many possible regulatory mechanisms acting at the level of transcription, it seems striking that the gene dose effects were so clearly detectable in gained areas. One hypothetical explanation may lie in the loss of controlled methylation in tumor cells (17–19). Thus, it may be possible that in chromosomes with increased DNA copy numbers two or more alleles could be demethylated simultaneously leading to a higher transcription level, whereas in chromosomes with losses the remaining allele could be partly methylated, turning off the process (20, 21). A recent report has documented a ploidy regulation of gene expression in yeast, but in this case all the genes were present in the same ratio (22), a situation that is not analogous to that of cancer cells, which show marked chromosomal aberrations, as well as gene dosage effects.

Several CGH studies of bladder cancer have shown that some chromosomal aberrations are common at certain stages of disease progression, often occurring in more than 1 of 3 tumors. In pTa tumors, these include 9p−, 9q−, 1q+, Y− (2, 6), and in pT1 tumors, 2q−,11p−, 11q−, 1q+, 5p+, 8q+, 17q+, and 20q+ (2–4, 6, 7). The pTa tumors studied here showed similar aberrations such as 9p− and 9q22-q33− and 9q− and Y−, respectively. Likewise, the two minimal invasive pT1 tumors showed aberrations that are commonly seen at that stage, and TCC 827 had a remarkable resemblance to the commonly seen pattern of losses and gains, such as 1q22–24 amplification (seen in both tumors), 11q14-q22 loss, the latter often linked to 17 q+ (both tumors), and 1q+ and 9p−, often linked to 20q+ and 11 q13+ (both tumors) (7–9). These observations indicate that the pairs of tumors used in this study exhibit chromosomal changes observed in many tumors, and therefore the findings could be of general importance for bladder cancer.

Considering that the mapping resolution of CGH is of about 20 megabases it is only possible to get a crude picture of chromosomal instability using this technique. Occasionally, we observed reduced transcript levels close to or inside regions with increased copy numbers. Analysis of these regions by positioning heterozygous microsatellites as close as possible to the locus showing reduced gene expression revealed loss of heterozygosity in several cases. It seems likely that multiple and different events occur along each chromosomal arm and that the use of cDNA microarrays for analysis of DNA copy number changes will reach a resolution that can resolve these changes, as has recently been proposed (2). The outlier data were not more frequent at the boundaries of the CGH aberrations. At present we do not know the mechanism behind chromosomal aneuploidy and cannot predict whether chromosomal gains will be transcribed to a larger extent than the two native alleles. A mechanism as genetic imprinting has an impact on the expression level in normal cells and is often reduced in tumors. However, the relation between imprinting and gain of chromosomal material is not known.

We regard it as a strength of this investigation that we were able to compare invasive tumors to benign tumors rather than to normal urothelium, as the tumors studied were biologically very close and probably may represent successive steps in the progression of bladder cancer. Despite the limited amount of fresh tissue available it was possible to apply three different state of the art methods. The observed correlation between DNA copy number and mRNA expression is remarkable when one considers that different pieces of the tumor biopsies were used for the different sets of experiments. This indicate that bladder tumors are relatively homogenous, a notion recently supported by CGH and LOH data that showed a remarkable similarity even between tumors and distant metastasis (10, 23).

In the few cases analyzed, mRNA and protein levels showed a striking correspondence although in some cases we found discrepancies that may be attributed to translational regulation, post-translational processing, protein degradation, or a combination of these. Some transcripts belong to undertranslated mRNA pools, which are associated with few translationally inactive ribosomes; these pools, however, seem to be rare (24). Protein degradation, for example, may be very important in the case of polypeptides with a short half-life (e.g. signaling proteins). A poor correlation between mRNA and protein levels was found in liver cells as determined by arrays and 2D-PAGE (25), and a moderate correlation was recently reported by Ideker *et al.* (26) in yeast.

Interestingly, our study revealed a much better correlation between gained chromosomal areas and increased mRNA levels than between loss of chromosomal areas and reduced mRNA levels. In general, the level of CGH change determined the ability to detect a change in transcript. One possible explanation could be that by losing one allele the change in mRNA level is not so dramatic as compared with gain of material, which can be rather unlimited and may lead to a severalfold increase in gene copy number resulting in a much higher impact on transcript level. The latter would be much easier to detect on the expression arrays as the cut-off point was placed at a 2-fold level so as not to be biased by noise on the array. Construction of arrays with a better signal to noise ratio may in the future allow detection of lesser than 2-fold alterations in transcript levels, a feature that may facilitate the analysis of the effect of loss of chromosomal areas on transcript levels.

In eleven cases we found a significant correlation between DNA copy number, mRNA expression, and protein level. Four of these proteins were encoded by genes located at a frequently amplified area in chromosome 17q. Whether DNA copy number is one of the mechanisms behind alteration of these eleven proteins is at present unknown and will have to be proved by other methods using a larger number of samples. One factor making such studies complicated is the large extent of protein modification that occurs after translation, requiring immunoidentification and/or mass spectrometry to correctly identify the proteins in the gels.

In conclusion, the results presented in this study exemplify the large body of knowledge that may be possible to gather in the future by combining state of the art techniques that follow the pathway from DNA to protein (26). Here, we used a traditional chromosomal CGH method, but in the future high resolution CGH based on microarrays with many thousand radiation hybrid-mapped genes will increase the resolution and information derived from these types of experiments (2). Combined with expression arrays analyzing transcripts derived from genes with known locations, and 2D gel analysis to obtain information at the post-translational level, a clearer and more developed understanding of the tumor genome will be forthcoming.

*Acknowledgments*—We thank Mie Madsen, Hanne Steen, Inge Lis Thorsen, Hans Lund, Vikolaj Ørntoft, and Lynn Bjerke for technical help and Thomas Gingeras, Christine Harrington, and Morten Østergaard for valuable discussions.

* This work was supported by grants from The Danish Cancer Society, the University of Aarhus, Aarhus County, Novo Nordic, the Danish Biotechnology Program, the Frenkels Foundation, the John and Birthe Meyer Foundation, and NCI, National Institutes of Health Grant CA47537. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: Dept. of Clinical Biochemistry, Molecular Diagnostic Laboratory, Aarhus University Hospital, Skejby, DK-8200 Aarhus N, Denmark. Tel.: 45-89495100/45-86156201 (private); Fax: 45-89496018; E-mail: orntoft@kba.sks.au.dk.

REFERENCES

1. Lengauer, C., Kinzler, K. W., and Vogelstein, B. (1998) Genetic instabilities in human cancers. *Nature* **17**, 643–649
2. Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D., and Brown, P. O. (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.* **23**, 41–46
3. de Cremoux, P., Martin, E. C., Vincent-Salomon, A., Dieras, V., Barbaroux, C., Liva, S., Pouillart, P., Sastre-Garau, X., and Magdelenat, H. (1999) Quantitative PCR analysis of c-erb B-2 (HER2/neu) gene amplification and comparison with p185(HER2/neu) protein expression in breast cancer drill biopsies. *Int. J. Cancer* **83**, 157–161
4. Brunger, P. P., Tamimi, Y., Shuuring, E., and Schalken, J. (1996) Expression of cyclin D1 and EMS1 in bladder tumors; relationship with chromosome 11q13 amplifications. *Oncogene* **12**, 1747–1753
5. Slavc, I., Ellenbogen, R., Jung, W. H., Vawter, G. F., Kretschmar, C., Grier, H., and Korf, B. R. (1990) *myc* gene amplification and expression in primary human neuroblastoma. *Cancer Res.* **50**, 1459–1463
6. Sauter, G., Carroll, P., Moch, H., Kallioniemi, A., Kerschmann, R., Narayan, P., Mihatsch, M. J., and Waldman, F. M. (1995) c-*myc* copy number gains in bladder cancer detected by fluorescence *in situ* hybridization. *Am J. Pathol.* **146**, 1131–1139
7. Richter, J., Jiang, F., Gorog, J. P., Sartorious, G., Egenter, C., Gasser, T. C., Moch, H., Mihatsch, M. J., and Sauter, G. (1997) Marked genetic differences between stage pTa and stage pT1 papillary bladder cancer detected by comparative genomic hybridization. *Cancer Res.* **57**, 2860–2864
8. Richter, J., Beffa, L., Wagner, U., Schraml, P., Gasser, T. C., Moch, H., Mihatsch, M. J., and Sauter, G. (1998) Patterns of chromosomal imbalances in advanced urinary bladder cancer detected by comparative genomic hybridization. *Am. J. Pathol.* **153**, 1615–1621
9. Bruch, J., Wohr, G., Hautmann, R., Mattfeldt, T., Bruderlein, S., Moller, P., Sauter, S., Hameister, H., Vogel, W., and Paiss, T. (1998) Chromosomal changes during progression of transitional cell carcinoma of the bladder and delineation of the amplified interval on chromosome arm 8q. *Genes Chromosomes Cancer* **23**, 167–174
10. Hovey, R. M., Chu, L., Balazs, M., De Vries, S., Moore, D., Sauter, G., Carroll, P. R., and Waldman, F. M. (1998) Genetic alterations in primary bladder cancers and their metastases. *Cancer Res.* **15**, 3555–3560
11. Simon, R., Burger, H., Brinkschmidt, C., Bocker, W., Hertle, L., and Terpe, H. J. (1998) Chromosomal aberrations associated with invasion in papillary superficial bladder cancer. *J. Pathol.* **185**, 345–351
12. Koo, S. H., Kwon, K. C., Ihm, C. H., Jeon, Y. M., Park, J. W., and Sut, C. K. (1999) Detection of genetic alterations in bladder tumors by comparative genomic hybridization and cytogenetic analysis. *Cancer Genet. Cytogenet.* **110**, 87–93
13. Wodicka, L., Dong, H., Mittmann, M., Ho, M. H., and Lockhart, D. J. (1997) Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **15**, 1359–1367
14. Christensen, M., Sunde, L., Bolund, L., and Ornfoft, T. F. (1999) Comparison of three methods of microsatellite detection. *Scand. J. Clin. Lab. Invest.* **59**, 167–177
15. Celis, J. E., Ostergaard, M., Basse, B., Celis, A., Lauridsen, J. B., Ratz, G. P., Andersen, I., Hein, B., Wolf, H., Ornfoft, T. F., and Rasmussen, H. H. (1996) Loss of adipocyte-type fatty acid binding protein and other protein biomarkers is associated with progression of human bladder transitional cell carcinomas. *Cancer Res.* **56**, 4782–4790
16. Celis, J. E., Ratz, G., Basse, B., Lauridsen, J. B., and Celis, A. (1994) In *Cell Biology: A Laboratory Handbook* (Celis, J. E., ed) Vol. 3, pp. 222–230, Academic Press, Orlando, FL
17. Ohlsson, R., Tycko, B., and Sapienza, C. (1998) Monoallelic expression: 'there can only be one'. *Trends Genet.* **14**, 435–438
18. Hollander, G. A., Zuklys, S., Morel, C., Mizoguchi, E., Mobisson, K., Simpson, S., Terhorst, C., Wishart, W., Golan, D. E., Bhan, A. K., and Burakoff, S. J. (1998) Monoallelic expression of the interleukin-2 locus. *Science* **279**, 2118–2121
19. Brannan, C. I., and Bartolomei, M. S. (1999) Mechanisms of genomic imprinting. *Curr. Opin. Genet. Dev.* **9**, 164–170
20. Ohlsson, R., Cui, H., He, L., Pfeifer, S., Malmikumpa, H., Jiang, S., Feinberg, A. P., and Hedborg, F. (1999) Mosaic allelic insulin-like growth factor 2 expression patterns reveal a link between Wilms' tumorigenesis and epigenetic heterogeneity. *Cancer Res.* **59**, 3889–3892
21. Cui, H., Hedborg, F., He, L., Nordenskjold, A., Sandstedt, B., Pfeifer-Ohlsson, S., and Ohlsson, R. (1997) Inactivation of H19, an imprinted and putative tumor repressor gene, is a preneoplastic event during Wilms' tumorigenesis. *Cancer Res.* **57**, 4469–4473
22. Gafitski, T., Saldanha, A. J., Styles, C. A., Lander, E. S., and Fink, G. R. (1999) Ploidy regulation of gene expression. *Science* **285**, 251–254
23. Tsao, J., Yatabe, Y., Marki, I. D., Hajyan, K., Jones, P. A., and Shibata, D. (2000) Bladder cancer genotype stability during clinical progression. *Genes Chromosomes Cancer* **29**, 26–32
24. Zong, Q., Schummer, M., Hood, L., and Morris, D. R. (1999) Messenger RNA translation state: the second dimension of high-throughput expression screening. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10632–10636
25. Anderson, L., and Seilhamer, J. (1997) Comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* **18**, 533–537
26. Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., Bumgarner, R., Goodlett, D. R., Aebersold, R., and Hood, L. (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **292**, 929–934

[CANCER RESEARCH 62, 6240–6245, November 1, 2002]



# Impact of DNA Amplification on Gene Expression Patterns in Breast Cancer[1,2]

**Elizabeth Hyman,[3] Päivikki Kauraniemi,[3] Sampsa Hautaniemi, Maija Wolf, Spyro Mousses, Ester Rozenblum, Markus Ringnér, Guido Sauter, Outi Monni, Abdel Elkahloun, Olli-P. Kallioniemi, and Anne Kallioniemi[4]**

*Howard Hughes Medical Institute-NIH Research Scholar, Bethesda, Maryland 20892 [E. H.]; Cancer Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, Maryland 20892 [E. H., P. K., S. H., M. W., S. M., E. R., M. R., A. E., O. K., A. K.]; Laboratory of Cancer Genetics, Institute of Medical Technology, University of Tampere and Tampere University Hospital, FIN-33520 Tampere, Finland [P. K., A. K.]; Signal Processing Laboratory, Tampere University of Technology, FIN-33101 Tampere, Finland [S. H.]; Institute of Pathology, University of Basel, CH-4003 Basel, Switzerland [G. S.]; and Biomedicum Biochip Center, Helsinki University Hospital, Biomedicum Helsinki, FIN-00014 Helsinki, Finland [O. M.]*

## ABSTRACT

**Genetic changes underlie tumor progression and may lead to cancer-specific expression of critical genes. Over 1100 publications have described the use of comparative genomic hybridization (CGH) to analyze the pattern of copy number alterations in cancer, but very few of the genes affected are known. Here, we performed high-resolution CGH analysis on cDNA microarrays in breast cancer and directly compared copy number and mRNA expression levels of 13,824 genes to quantitate the impact of genomic changes on gene expression. We identified and mapped the boundaries of 24 independent amplicons, ranging in size from 0.2 to 12 Mb. Throughout the genome, both high- and low-level copy number changes had a substantial impact on gene expression, with 44% of the highly amplified genes showing overexpression and 10.5% of the highly overexpressed genes being amplified. Statistical analysis with random permutation tests identified 270 genes whose expression levels across 14 samples were systematically attributable to gene amplification. These included most previously described amplified genes in breast cancer and many novel targets for genomic alterations, including the *HOXB7* gene, the presence of which in a novel amplicon at 17q21.3 was validated in 10.2% of primary breast cancers and associated with poor patient prognosis. In conclusion, CGH on cDNA microarrays revealed hundreds of novel genes whose overexpression is attributable to gene amplification. These genes may provide insights to the clonal evolution and progression of breast cancer and highlight promising therapeutic targets.**

## INTRODUCTION

Gene expression patterns revealed by cDNA microarrays have facilitated classification of cancers into biologically distinct categories, some of which may explain the clinical behavior of the tumors (1–6). Despite this progress in diagnostic classification, the molecular mechanisms underlying gene expression patterns in cancer have remained elusive, and the utility of gene expression profiling in the identification of specific therapeutic targets remains limited.

Accumulation of genetic defects is thought to underlie the clonal evolution of cancer. Identification of the genes that mediate the effects of genetic changes may be important by highlighting transcripts that are actively involved in tumor progression. Such transcripts and their encoded proteins would be ideal targets for anticancer therapies, as demonstrated by the clinical success of new therapies against amplified oncogenes, such as *ERBB2* and *EGFR* (7, 8), in breast cancer and other solid tumors. Besides amplifications of known oncogenes, over 20 recurrent regions of DNA amplification have been mapped in breast cancer by CGH[5] (9, 10). However, these amplicons are often large and poorly defined, and their impact on gene expression remains unknown.

We hypothesized that genome-wide identification of those gene expression changes that are attributable to underlying gene copy number alterations would highlight transcripts that are actively involved in the causation or maintenance of the malignant phenotype. To identify such transcripts, we applied a combination of cDNA and CGH microarrays to: (*a*) determine the global impact that gene copy number variation plays in breast cancer development and progression; and (*b*) identify and characterize those genes whose mRNA expres-

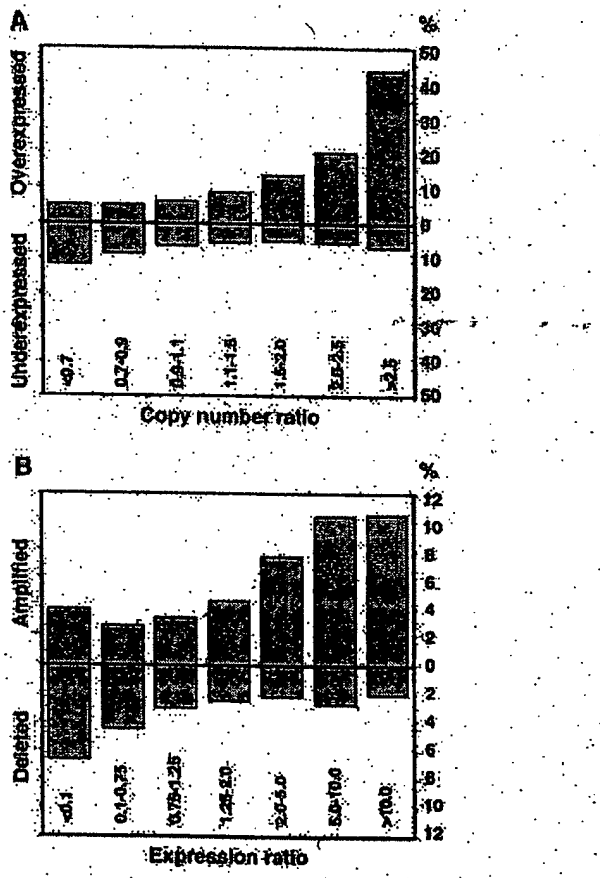



Fig. 1. Impact of gene copy number on global gene expression levels. *A*, percentage of over- and underexpressed genes (*Y axis*) according to copy number ratios (*X axis*). Threshold values used for over- and underexpression were >2.184 (global upper 7% of the cDNA ratios) and <0.4826 (global lower 7% of the expression ratios). *B*, percentage of amplified and deleted genes according to expression ratios. Threshold values for amplification and deletion were >1.5 and <0.7.

Received 5/29/02; accepted 8/28/02.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported in part by the Academy of Finland, Emil Aaltonen Foundation, the Finnish Cancer Society, the Pirkanmaa Cancer Society, the Pirkanmaa Cultural Foundation, the Finnish Breast Cancer Group, the Foundation for the Development of Laboratory Medicine, the Medical Research Fund of the Tampere University Hospital, the Foundation for Commercial and Technical Sciences, and the Swedish Research Council.

[2] Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org).

[3] Contributed equally to this work.

[4] To whom requests for reprints should be addressed, at Laboratory of Cancer Genetics, Institute of Medical Technology, Lenkkeilijankatu 6, FIN-33520 Tampere, Finland. Phone: 358-3247-4125; Fax: 358-3247-4168; E-mail: anne.kallioniemi@uta.fi.

[5] The abbreviations used are: CGH, comparative genomic hybridization; FISH, fluorescence *in situ* hybridization; RT-PCR, reverse transcription-PCR.

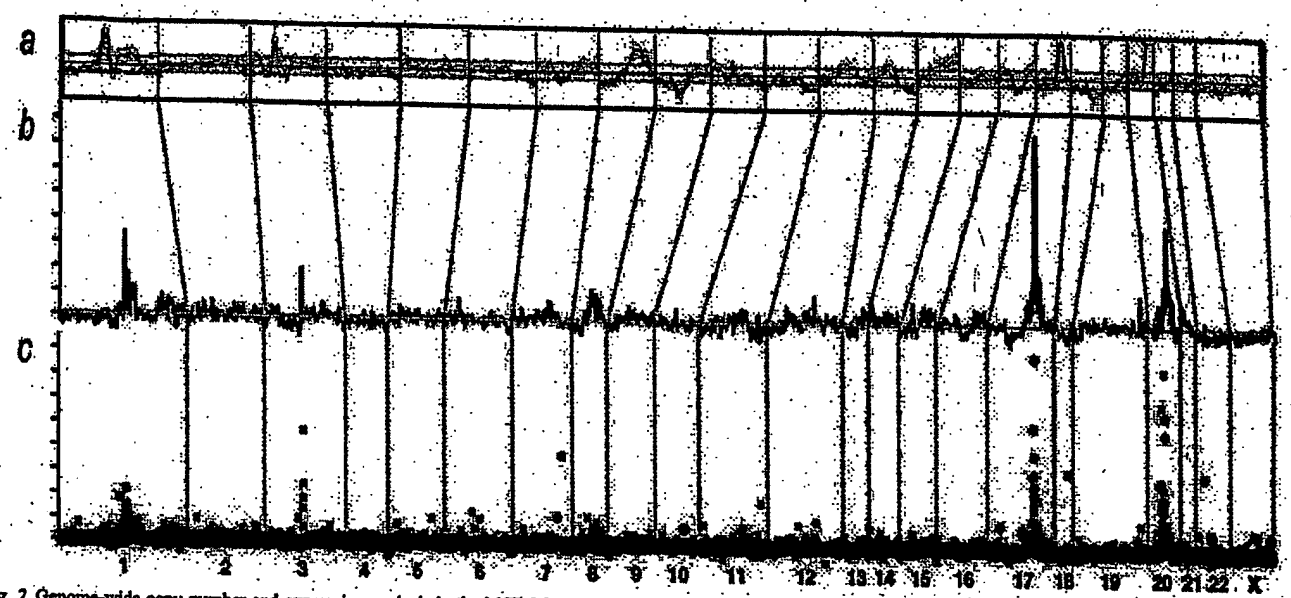



Fig. 2. Genome-wide copy number and expression analysis in the MCF-7 breast cancer cell line. *A*, chromosomal CGH analysis of MCF-7. The copy number ratio profile (*blue line*) across the entire genome from 1p telomere to Xq telomere is shown along with ±1 SD (*orange lines*). The *black horizontal line* indicates a ratio of 1.0; *red line*, a ratio of 0.8; and *green line*, a ratio of 1.2. *B–C*, genome-wide copy number analysis in MCF-7 by CGH on cDNA microarray. The copy number ratios were plotted as a function of the position of the cDNA clones along the human genome. In *B*, individual data points are connected with a line, and a moving median of 10 adjacent clones is shown. *Red horizontal line*, the copy number ratio of 1.0. In *C*, individual data points are labeled by color coding according to cDNA expression ratios. The *bright red dots* indicate the upper 2%, and *dark red dots*, the next 5% of the expression ratios in MCF-7 cells (overexpressed genes); *bright green dots* indicate the lowest 2%, and *dark green dots*, the next 5% of the expression ratios (underexpressed genes); the rest of the observations are shown with *black crosses*. The chromosome numbers are shown at the *bottom* of the figure, and chromosome boundaries are indicated with a *dashed line*.

sion is most significantly associated with amplification of the corresponding genomic template.

## MATERIALS AND METHODS

**Breast Cancer Cell Lines.** Fourteen breast cancer cell lines (BT-20, BT-474, HCC1428, Hs578t, MCF7, MDA-361, MDA-436, MDA-453, MDA-468, SKBR-3, T-47D, UACC812, ZR-75-1, and ZR-75-30) were obtained from the American Type Culture Collection (Manassas, VA). Cells were grown under recommended culture conditions. Genomic DNA and mRNA were isolated using standard protocols.

**Copy Number and Expression Analyses by cDNA Microarrays.** The preparation and printing of the 13,824 cDNA clones on glass slides were performed as described (11–13). Of these clones, 244 represented uncharacterized expressed sequence tags, and the remainder corresponded to known genes. CGH experiments on cDNA microarrays were done as described (14, 15). Briefly, 20 $\mu$g of genomic DNA from breast cancer cell lines and normal human WBCs were digested for 14–18 h with *Alu*I and *Rsa*I (Life Technologies, Inc., Rockville, MD) and purified by phenol/chloroform extraction. Six $\mu$g of digested cell line DNAs were labeled with Cy3-dUTP (Amersham Pharmacia) and normal DNA with Cy5-dUTP (Amersham Pharmacia) using the Bioprime Labeling kit (Life Technologies, Inc.). Hybridization (14, 15) and posthybridization washes (13) were done as described. For the expression analyses, a standard reference (Universal Human Reference RNA; Stratagene, La Jolla, CA) was used in all experiments. Forty $\mu$g of reference RNA were labeled with Cy3-dUTP and 3.5 $\mu$g of test mRNA with Cy5-dUTP, and the labeled cDNAs were hybridized on microarrays as described (13, 15). For both microarray analyses, a laser confocal scanner (Agilent Technologies, Palo Alto, CA) was used to measure the fluorescence intensities at the target locations using the DEARRAY software (16). After background subtraction, average intensities at each clone in the test hybridization were divided by the average intensity of the corresponding clone in the control hybridization. For the copy number analysis, the ratios were normalized on the basis of the distribution of ratios of all targets on the array and for the expression analysis on the basis of 88 housekeeping genes, which were spotted four times onto the array. Low quality measurements (*i.e.*, copy number data with mean reference intensity <100 fluorescent units, and expression data with both test and reference intensity <100 fluorescent units and/or with spot size <50 units) were excluded from the analysis and were treated as missing values. The distributions of fluorescence ratios were used to define cutpoints for increased/decreased copy number. Genes with CGH ratio >1.43 (representing the upper 5% of the CGH ratios across all experiments) were considered to be amplified, and genes with ratio <0.73 (representing the lower 5%) were considered to be deleted.

**Statistical Analysis of CGH and cDNA Microarray Data.** To evaluate the influence of copy number alterations on gene expression, we applied the following statistical approach. CGH and cDNA calibrated intensity ratios were log-transformed and normalized using median centering of the values in each cell line. Furthermore, cDNA ratios for each gene across all 14 cell lines were median centered. For each gene, the CGH data were represented by a vector that was labeled 1 for amplification (ratio, >1.43) and 0 for no amplification. Amplification was correlated with gene expression using the signal-to-noise statistics (1). We calculated a weight, $w_g$, for each gene as follows:

$$w_g = \frac{m_{g1} - m_{g0}}{\sigma_{g1} + \sigma_{g0}}$$

where $m_{g1}$, $\sigma_{g1}$ and $m_{g0}$, $\sigma_{g0}$ denote the means and SDs for the expression levels for amplified and nonamplified cell lines, respectively. To assess the statistical significance of each weight, we performed 10,000 random permutations of the label vector. The probability that a gene had a larger or equal weight by random permutation than the original weight was denoted by $\alpha$. A low $\alpha$ (<0.05) indicates a strong association between gene expression and amplification.

**Genomic Localization of cDNA Clones and Amplicon Mapping.** Each cDNA clone on the microarray was assigned to a Unigene cluster using the Unigene Build 141.[6] A database of genomic sequence alignment information for mRNA sequences was created from the August 2001 freeze of the University of California Santa Cruz's GoldenPath database.[7] The chromosome and bp positions for each cDNA clone were then retrieved by relating these data sets. Amplicons were defined as a CGH copy number ratio >2.0 in at least two adjacent clones in two or more cell lines or a CGH ratio >2.0 in at least three adjacent clones in a single cell line. The amplicon start and end positions were

[6] Internet address: http://research.nhgri.nih.gov/microarray/downloadable_cdna.html.
[7] Internet address: www.genome.ucsc.edu.

Table 1 *Summary of independent amplicons in 14 breast cancer cell lines by CGH microarray*

| Location | Start (Mb) | End (Mb) | Size (Mb) |
|---|---|---|---|
| 1p13 | 132.79 | 132.94 | 0.2 |
| 1q21 | 173.92 | 177.25 | 3.3 |
| 1q22 | 179.28 | 179.57 | 0.3 |
| 3p14 | 71.94 | 74.66 | 2.7 |
| 7p12.1–7p11.2 | 55.62 | 60.95 | 5.3 |
| 7q31 | 125.73 | 130.96 | 5.2 |
| 7q32 | 140.01 | 140.68 | 0.7 |
| 8q21.11–8q21.13 | 86.45 | 92.46 | 6.0 |
| 8q21.3 | 98.45 | 103.05 | 4.6 |
| 8q23.3–8q24.14 | 129.88 | 142.15 | 12.3 |
| 8q24.22 | 151.21 | 152.16 | 1.0 |
| 9p13 | 38.65 | 39.25 | 0.6 |
| 13q22–q31 | 77.15 | 81.38 | 4.2 |
| 16q22 | 86.70 | 87.62 | 0.9 |
| 17q11 | 29.30 | 30.85 | 1.6 |
| 17q12–q21.2 | 39.79 | 42.80 | 3.0 |
| 17q21.32–q21.33 | 52.47 | 55.80 | 3.3 |
| 17q22–q23.3 | 63.81 | 69.70 | 5.9 |
| 17q23.3–q24.3 | 69.93 | 74.99 | 5.1 |
| 19q13 | 40.63 | 41.40 | 0.8 |
| 20q11.22 | 34.59 | 35.85 | 1.3 |
| 20q13.12 | 44.00 | 45.62 | 1.6 |
| 20q13.12–q13.13 | 46.45 | 49.43 | 3.0 |
| 20q13.2–q13.32 | 51.32 | 59.12 | 7.8 |

extended to include neighboring nonamplified clones (ratio, <1.5). The amplicon size determination was partially dependent on local clone density.

**FISH.** Dual-color interphase FISH to breast cancer cell lines was done as described (17). Bacterial artificial chromosome clone RP11-361K8 was labeled with SpectrumOrange (Vysis, Downers Grove, IL), and SpectrumOrange-labeled probe for *EGFR* was obtained from Vysis. SpectrumGreen-labeled chromosome 7 and 17 centromere probes (Vysis) were used as a reference. A tissue microarray containing 612 formalin-fixed, paraffin-embedded primary breast cancers (17) was applied in FISH analyses as described (18). The use of these specimens was approved by the Ethics Committee of the University of Basel and by the NIH. Specimens containing a 2-fold or higher increase in the number of test probe signals, as compared with corresponding centromere signals, in at least 10% of the tumor cells were considered to be amplified. Survival analysis was performed using the Kaplan-Meier method and the log-rank test.

**RT-PCR.** The *HOXB7* expression level was determined relative to *GAPDH*. Reverse transcription and PCR amplification were performed using Access RT-PCR System (Promega Corp., Madison, WI) with 10 ng of mRNA as a template. *HOXB7* primers were 5′-GAGCAGAGGGACTCGGACTT-3′ and 5′-GCGTCAGGTAGCGATTGTAG-3′.

## RESULTS

**Global Effect of Copy Number on Gene Expression.** 13,824 arrayed cDNA clones were applied for analysis of gene expression and gene copy number (CGH microarrays) in 14 breast cancer cell lines. The results illustrate a considerable influence of copy number on gene expression patterns. Up to 44% of the highly amplified transcripts (CGH ratio, >2.5) were overexpressed (*i.e.*, belonged to the global upper 7% of expression ratios), compared with only 6% for genes with normal copy number levels (Fig. 1*A*). Conversely, 10.5% of the transcripts with high-level expression (cDNA ratio, >10) showed increased copy number (Fig. 1*B*). Low-level copy number increases and decreases were also associated with similar, although less dramatic, outcomes on gene expression (Fig. 1).

**Identification of Distinct Breast Cancer Amplicons.** Base-pair locations obtained for 11,994 cDNAs (86.8%) were used to plot copy number changes as a function of genomic position (Fig. 2, Supplement Fig. A). The average spacing of clones throughout the genome was 267 kb. This high-resolution mapping identified 24 independent breast cancer amplicons, spanning from 0.2 to 12 Mb of DNA (Table 1). Several amplification sites detected previously by chromosomal CGH were validated, with 1q21, 17q12–q21.2, 17q22–q23, 20q13.1, and 20q13.2 regions being most commonly amplified. Furthermore, the boundaries of these amplicons were precisely delineated. In addition, novel amplicons were identified at 9p13 (38.65–39.25 Mb), and 17q21.3 (52.47–55.80 Mb).

**Direct Identification of Putative Amplification Target Genes.** The cDNA/CGH microarray technique enables the direct correlation of copy number and expression data on a gene-by-gene basis throughout the genome. We directly annotated high-resolution CGH plots with gene expression data using color coding. Fig. 2*C* shows that most of the amplified genes in the MCF-7 breast cancer cell line at 1p13, 17q22–q23, and 20q13 were highly overexpressed. A view of chromosome 7 in the MDA-468 cell line implicates *EGFR* as the most highly overexpressed and amplified gene at 7p11–p12 (Fig. 3*A*). In BT-474, the two known amplicons at 17q12 and 17q22–q23 contained numerous highly overexpressed genes (Fig. 3*B*). In addition, several genes, including the homeobox genes *HOXB2* and *HOXB7*, were highly amplified in a previously undescribed independent amplicon at 17q21.3. *HOXB7* was systematically amplified (as validated by FISH, Fig. 3*B*, *inset*) as well as overexpressed (as verified by RT-PCR, data not shown) in BT-474, UACC812, and ZR-75-30 cells. Furthermore, this novel


Fig. 3. Annotation of gene expression data on CGH microarray profiles. *A*, genes in the 7p11-p12 amplicon in the MDA-468 cell line are highly expressed (*red dots*) and include the *EGFR* oncogene. *B*, several genes in the 17q12, 17q21.3, and 17q23 amplicons in the BT-474 breast cancer cell line are highly overexpressed (*red*) and include the *HOXB7* gene. The data labels and color coding are as indicated for Fig. 2*C*. *Insets* show chromosomal CGH profiles for the corresponding chromosomes and validation of the increased copy number by interphase FISH using *EGFR* (*red*) and chromosome 7 centromere probe (*green*) to MDA-468 (*A*) and *HOXB7*-specific probe (*red*) and chromosome 17 centromere (*green*) to BT-474 cells (*B*).

level copy number increase. Low-level copy number gains and losses also had a significant influence on expression levels of genes in the regions affected, but these effects were more subtle on a gene-by-gene basis than those of high-level amplifications. However, the impact of low-level gains on the dysregulation of gene expression patterns in cancer may be equally important if not more important than that of high-level amplifications. Aneuploidy and low-level gains and losses of chromosomal arms represent the most common types of genetic alterations in breast and other cancers and, therefore, have an influence on many genes. Our results in breast cancer extend the recent studies on the impact of aneuploidy on global gene expression patterns in yeast cells, acute myeloid leukemia, and a prostate cancer model system (22–24).

The CGH microarray analysis identified 24 independent breast cancer amplicons. We defined the precise boundaries for many amplicons detected previously by chromosomal CGH (9, 10, 25, 26) and also discovered novel amplicons that had not been detected previously, presumably because of their small size (only 1–2 Mb) or close proximity to other larger amplicons. One of these novel amplicons involved the homeobox gene region at 17q21.3 and led to the overexpression of the *HOXB7* and *HOXB2* genes. The homeodomain transcription factors are known to be key regulators of embryonic development and have been occasionally reported to undergo aberrant expression in cancer (27, 28). *HOXB7* transfection induced cell proliferation in melanoma, breast, and ovarian cancer cells and increased tumorigenicity and angiogenesis in breast cancer (29–32). The present results imply that gene amplification may be a prominent mechanism for overexpressing *HOXB7* in breast cancer and suggest that *HOXB7* contributes to tumor progression and confers an aggressive disease phenotype in breast cancer. This view is supported by our finding of amplification of *HOXB7* in 10% of 363 primary breast cancers, as well as an association of amplification with poor prognosis of the patients.

We carried out a systematic search to identify genes whose expression levels across all 14 cell lines were attributable to amplification status. Statistical analysis revealed 270 such genes (representing ~2% of all genes on the array), including not only previously described amplified genes, such as *HER-2*, *MYC*, *EGFR*, ribosomal protein s6 kinase, and *AIB3*, but also numerous novel genes such as *NRAS-related gene* (1p13), *syndecan-2* (8q22), and *bone morphogenic protein* (20q13.1), whose activation by amplification may similarly promote breast cancer progression. Most of the 270 genes have not been implicated previously in breast cancer development and suggest novel pathogenetic mechanisms. Although we would not expect all of them to be causally involved, it is intriguing that 84% of the genes with associated functional information were implicated in apoptosis, cell proliferation, signal transduction, transcription, or other cellular processes that could directly imply a possible role in cancer progression. Therefore, a detailed characterization of these genes may provide biological insights to breast cancer progression and might lead to the development of novel therapeutic strategies.

In summary, we demonstrate application of cDNA microarrays to the analysis of both copy number and expression levels of over 12,000 transcripts throughout the breast cancer genome, roughly once every 267 kb. This analysis provided: (*a*) evidence of a prominent global influence of copy number changes on gene expression levels; (*b*) a high-resolution map of 24 independent amplicons in breast cancer; and (*c*) identification of a set of 270 genes, the overexpression of which was statistically attributable to gene amplification. Characterization of a novel amplicon at 17q21.3 implicated amplification and overexpression of the *HOXB7* gene in breast cancer, including a clinical association between *HOXB7* amplification and poor patient prognosis. Overall, our results illustrate how the identification of genes activated by gene amplification provides a powerful approach to highlight genes with an important role in cancer as well as to prioritize and validate putative targets for therapy development.

## REFERENCES


1. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science (Wash. DC), *286:* 531–537, 1999.
2. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature (Lond.), *403:* 503–511, 2000.
3. Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., *et al.* Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature (Lond.), *406:* 536–540, 2000.
4. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., *et al.* Molecular portraits of human breast tumours. Nature (Lond.), *406:* 747–752, 2000.
5. Dhanasekaran, S. M., Barrette, T. R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K. J., Rubin, M. A., and Chinnaiyan, A. M. Delineation of prognostic biomarkers in prostate cancer. Nature (Lond.), *412:* 822–826, 2001.
6. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc. Natl. Acad. Sci. USA, *98:* 10869–10874, 2001.
7. Ross, J. S., and Fletcher, J. A. The *HER-2/neu* oncogene: prognostic factor, predictive factor and target for therapy. Semin. Cancer Biol., *9:* 125–138, 1999.
8. Arteaga, C. L. The epidermal growth factor receptor: from mutant oncogene in nonhuman cancers to therapeutic target in human neoplasia. J. Clin. Oncol., *19:* 32–40, 2001.
9. Knuutila, S., Bjorkqvist, A. M., Autio, K., Tarkkanen, M., Wolf, M., Monni, O., Szymanska, J., Larramendy, M. L., Tapper, J., Pere, H., El-Rifai, W., *et al.* DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies. Am. J. Pathol., *152:* 1107–1123, 1998.
10. Knuutila S., Autio K., and Aalto Y. Online access to CGH data of DNA sequence copy number changes. Am. J. Pathol., *157:* 689, 2000.
11. DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A., and Trent, J. M. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat. Genet., *14:* 457–460, 1996.
12. Shalon, D., Smith, S. J., and Brown, P. O. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. Genome Res., *6:* 639–645, 1996.
13. Mousses, S., Bittner, M. L., Chen, Y., Dougherty, E. R., Baxevanis, A., Meltzer, P. S., and Trent, J. M. Gene expression analysis by cDNA microarrays. *In:* F. J. Livesey and S. P. Hunt (eds.), Functional Genomics, pp. 113–137. Oxford: Oxford University Press, 2000.
14. Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D., and Brown, P. O. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat. Genet., *23:* 41–46, 1999.
15. Monni, O., Bärlund, M., Mousses, S., Kononen, J., Sauter, G., Heiskanen, M., Paavola, P., Avela, K., Chen, Y., Bittner, M. L., and Kallioniemi, A. Comprehensive copy number and gene expression profiling of the 17q23 amplicon in human breast cancer. Proc. Natl. Acad. Sci. USA, *98:* 5711–5716, 2001.
16. Chen, Y., Dougherty, E. R., and Bittner, M. L. Ratio-based decisions and the quantitative analysis of cDNA microarray images. J. Biomed. Optics, *2:* 364–374, 1997.
17. Bärlund, M., Forozan, F., Kononen, J., Bubendorf, L., Chen, Y., Bittner, M. L., Torhorst, J., Haas, P., Bucher, C., Sauter, G., *et al.* Detecting activation of ribosomal protein S6 kinase by complementary DNA and tissue microarray analysis. J. Natl. Cancer Inst., *92:* 1252–1259, 2000.
18. Andersen, C. L., Hostetter, G., Grigoryan, A., Sauter, G., and Kallioniemi, A. Improved procedure for fluorescence *in situ* hybridization on tissue microarrays. Cytometry, *45:* 83–86, 2001.
19. Kauraniemi, P., Bärlund, M., Monni, O., and Kallioniemi, A. New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays. Cancer Res., *61:* 8235–8240, 2001.
20. Clark, J., Edwards, S., John, M., Flohr, P., Gordon, T., Maillard, K., Giddings, I., Brown, C., Bagherzadeh, A., Campbell, C., Shipley, J., Wooster, R., and Cooper, C. S. Identification of amplified and expressed genes in breast cancer by comparative hybridization onto microarrays of randomly selected cDNA clones. Genes Chromosomes Cancer, *34:* 104–114, 2002.
21. Varis, A., Wolf, M., Monni, O., Vakkari, M. L., Kokkola, A., Moskaluk, C., Frierson, H., Powell, S. M., Knuutila, S., Kallioniemi, A., and El-Rifai, W. Targets of gene amplification and overexpression at 17q in gastric cancer. Cancer Res., *62:* 2625–2629, 2002.
22. Hughes, T. R., Roberts, C. J., Dai, H., Jones, A. R., Meyer, M. R., Slade, D., Burchard, J., Dow, S., Ward, T. R., Kidd, M. J., Friend, S. H., and Marton M. J.

Widespread aneuploidy revealed by DNA microarray expression profiling. Nat. Genet., *25:* 333–337, 2000.

23. Virtaneva, K., Wright, F. A., Tanner, S. M., Yuan, B., Lemon, W. J., Caligiuri, M. A., Bloomfield, C. D., de La Chapelle, A., and Krahe, R. Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. Proc. Natl. Acad. Sci. USA, *98:* 1124–1129, 2001.
24. Phillips, J. L., Hayward, S. W., Wang, Y., Vasselli, J., Pavlovich, C., Padilla-Nash, H., Pezullo, J. R., Ghadimi, B. M., Grossfeld, G. D., Rivera, A., Linehan, W. M., Cunha, G. R., and Ried, T. The consequences of chromosomal aneuploidy on gene expression profiles in a cell line model for prostate carcinogenesis. Cancer Res., *61:* 8143–8149, 2001.
25. Bärlund, M., Tirkkonen, M., Forozan, F., Tanner, M. M., Kallioniemi, O. P., and Kallioniemi, A. Increased copy number at 17q22–q24 by CGH in breast cancer is due to high-level amplification of two separate regions. Genes Chromosomes Cancer, *20:* 372–376, 1997.
26. Tanner, M. M., Tirkkonen, M., Kallioniemi, A., Isola, J., Kuukasjärvi, T., Collins, C., Kowbel, D., Guan, X. Y., Trent, J., Gray, J. W., Meltzer, P., and Kallioniemi O. P. Independent amplification and frequent co-amplification of three nonsyntenic regions on the long arm of chromosome 20 in human breast cancer. Cancer Res., *56:* 3441–3445, 1996.
27. Cillo, C., Faiella, A., Cantile, M., and Boncinelli, E. Homeobox genes and cancer. Exp. Cell Res., *248:* 1–9, 1999.
28. Cillo, C., Cantile, M., Faiella, A., and Boncinelli, E. Homeobox genes in normal and malignant cells. J. Cell. Physiol., *188:* 161–169, 2001.
29. Care, A., Silvani, A., Meccia, E., Mattia, G., Stoppacciaro, A., Parmiani, G., Peschle, C., and Colombo, M. P. HOXB7 constitutively activates basic fibroblast growth factor in melanomas. Mol. Cell. Biol., *16:* 4842–4851, 1996.
30. Care, A., Silvani, A., Meccia, E., Mattia, G., Peschle, C., and Colombo, M. P. Transduction of the SkBr3 breast carcinoma cell line with the *HOXB7* gene induces bFGF expression, increases cell proliferation and reduces growth factor dependence. Oncogene, *16:* 3285–3289, 1998.
31. Care, A., Felicetti, F., Meccia, E., Bottero, L., Parenza, M., Stoppacciaro, A., Peschle, C., and Colombo, M. P. HOXB7: a key factor for tumor-associated angiogenic switch. Cancer Res., *61:* 6532–6539, 2001.
32. Naora, H., Yang, Y. Q., Montz, F. J., Seidman, J. D., Kurman, R. J., and Roden, R. B. A serologically identified tumor antigen encoded by a homeobox gene promotes growth of ovarian epithelial cells. Proc. Natl. Acad. Sci. USA, *98:* 4060–4065, 2001.

# Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors

Jonathan R. Pollack*†‡, Therese Sørlie§, Charles M. Perou¶, Christian A. Rees‖**, Stefanie S. Jeffrey††, Per E. Lonning‡‡, Robert Tibshirani§§, David Botstein‖, Anne-Lise Børresen-Dale§, and Patrick O. Brown†¶¶

Departments of *Pathology, ‖Genetics, ††Surgery, §§Health Research and Policy, and ¶¶Biochemistry, and †Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305; §Department of Genetics, Norwegian Radium Hospital, Montebello, N-0310 Oslo, Norway; ‡‡Department of Medicine (Oncology), Haukeland University Hospital, N-5021 Bergen, Norway; and ¶Department of Genetics and Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599

Contributed by Patrick O. Brown, August 6, 2002

**Genomic DNA copy number alterations are key genetic events in the development and progression of human cancers. Here we report a genome-wide microarray comparative genomic hybridization (array CGH) analysis of DNA copy number variation in a series of primary human breast tumors. We have profiled DNA copy number alteration across 6,691 mapped human genes, in 44 predominantly advanced, primary breast tumors and 10 breast cancer cell lines. While the overall patterns of DNA amplification and deletion corroborate previous cytogenetic studies, the high-resolution (gene-by-gene) mapping of amplicon boundaries and the quantitative analysis of amplicon shape provide significant improvement in the localization of candidate oncogenes. Parallel microarray measurements of mRNA levels reveal the remarkable degree to which variation in gene copy number contributes to variation in gene expression in tumor cells. Specifically, we find that 62% of highly amplified genes show moderately or highly elevated expression, that DNA copy number influences gene expression across a wide range of DNA copy number alterations (deletion, low-, mid- and high-level amplification), that on average, a 2-fold change in DNA copy number is associated with a corresponding 1.5-fold change in mRNA levels, and that overall, at least 12% of all the variation in gene expression among the breast tumors is directly attributable to underlying variation in gene copy number. These findings provide evidence that widespread DNA copy number alteration can lead directly to global deregulation of gene expression, which may contribute to the development or progression of cancer.**

Conventional cytogenetic techniques, including comparative genomic hybridization (CGH) (1), have led to the identification of a number of recurrent regions of DNA copy number alteration in breast cancer cell lines and tumors (2–4). While some of these regions contain known or candidate oncogenes [e.g., FGFR1 (8p11), MYC (8q24), CCND1 (11q13), ERBB2 (17q12), and ZNF217 (20q13)] and tumor suppressor genes [RB1 (13q14) and TP53 (17p13)], the relevant gene(s) within other regions (e.g., gain of 1q, 8q22, and 17q22–24, and loss of 8p) remain to be identified. A high-resolution genome-wide map, delineating the boundaries of DNA copy number alterations in tumors, should facilitate the localization and identification of oncogenes and tumor suppressor genes in breast cancer. In this study, we have created such a map, using array-based CGH (5–7) to profile DNA copy number alteration in a series of breast cancer cell lines and primary tumors.

An unresolved question is the extent to which the widespread DNA copy number changes that we and others have identified in breast tumors alter expression of genes within involved regions. Because we had measured mRNA levels in parallel in the same samples (8), using the same DNA microarrays, we had an opportunity to explore on a genomic scale the relationship between DNA copy number changes and gene expression. From this analysis, we have identified a significant impact of widespread DNA copy number alteration on the transcriptional programs of breast tumors.

## Materials and Methods

**Tumors and Cell Lines.** Primary breast tumors were predominantly large (>3 cm), intermediate-grade, infiltrating ductal carcinomas, with more than 50% being lymph node positive. The fraction of tumor cells within specimens averaged at least 50%. Details of individual tumors have been published (8, 9), and are summarized in Table 1, which is published as supporting information on the PNAS web site, www.pnas.org. Breast cancer cell lines were obtained from the American Type Culture Collection. Genomic DNA was isolated either using Qiagen genomic DNA columns, or by phenol/chloroform extraction followed by ethanol precipitation.

**DNA Labeling and Microarray Hybridizations.** Genomic DNA labeling and hybridizations were performed essentially as described in Pollack *et al.* (7), with slight modifications. Two micrograms of DNA was labeled in a total volume of 50 microliters and the volumes of all reagents were adjusted accordingly. "Test" DNA (from tumors and cell lines) was fluorescently labeled (Cy5) and hybridized to a human cDNA microarray containing 6,691 different mapped human genes (i.e., UniGene clusters). The "reference" (labeled with Cy3) for each hybridization was normal female leukocyte DNA from a single donor. The fabrication of cDNA microarrays and the labeling and hybridization of mRNA samples have been described (8).

**Data Analysis and Map Positions.** Hybridized arrays were scanned on a GenePix scanner (Axon Instruments, Foster City, CA), and fluorescence ratios (test/reference) calculated using SCANALYZE software (available at http://rana.lbl.gov). Fluorescence ratios were normalized for each array by setting the average log fluorescence ratio for all array elements equal to 0. Measurements with fluorescence intensities more than 20% above background were considered reliable. DNA copy number profiles that deviated significantly from background ratios measured in normal genomic DNA control hybridizations were interpreted as evidence of real DNA copy number alteration (see *Estimating Significance of Altered Fluorescence Ratios* in the supporting information). When indicated, DNA copy number profiles are displayed as a moving average (symmetric 5-nearest neighbors). Map positions for arrayed human cDNAs were assigned by

Abbreviation: CGH, comparative genomic hybridization.

†To whom reprint requests should be addressed at: Department of Pathology, Stanford University School of Medicine, CCSR Building, Room 3245A, 269 Campus Drive, Stanford, CA 94305-5176. E-mail: pollack1@stanford.edu.

**Present address: Zyomyx Inc., Hayward, CA 94545.

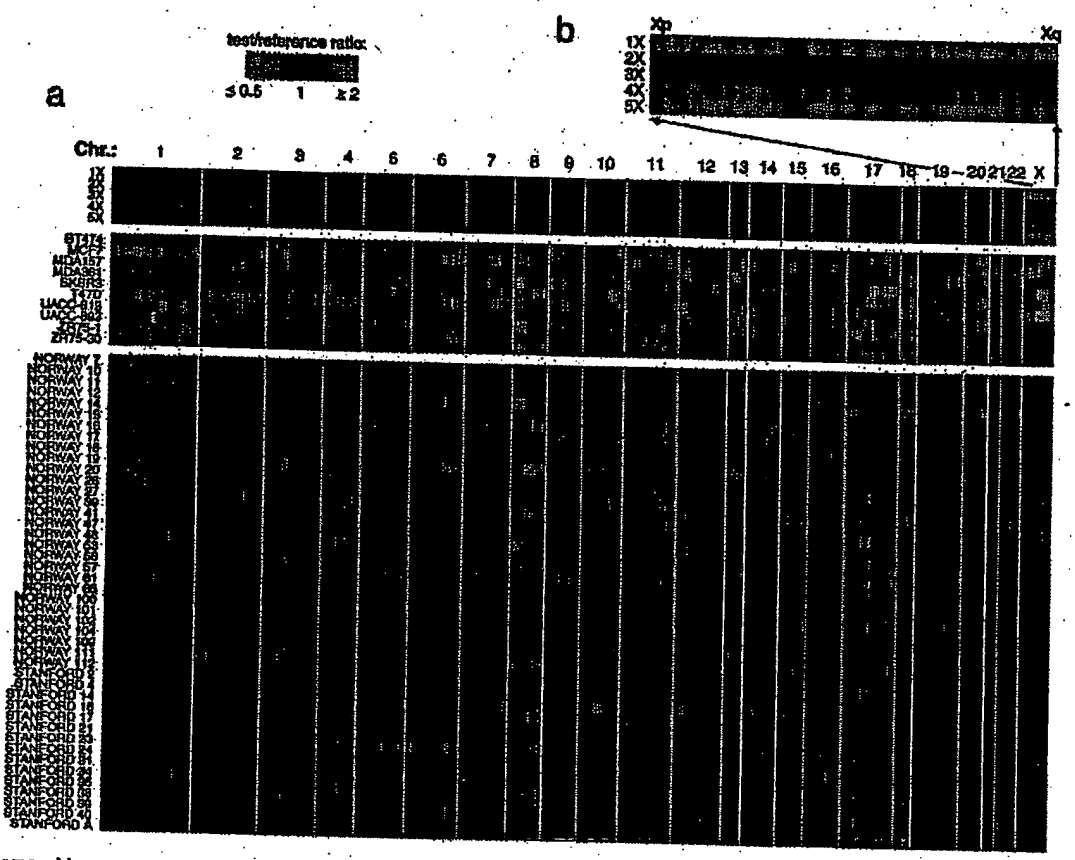



Fig. 1. Genome-wide measurement of DNA copy number alteration by array CGH. (*a*) DNA copy number profiles are illustrated for cell lines containing different numbers of X chromosomes, for breast cancer cell lines, and for breast tumors. Each row represents a different cell line or tumor, and each column represents one of 6,691 different mapped human genes present on the microarray, ordered by genome map position from 1pter through Xqter. Moving average (symmetric 5-nearest neighbors) fluorescence ratios (test/reference) are depicted using a $\log_2$-based pseudocolor scale (indicated), such that red luminescence reflects fold-amplification, green luminescence reflects fold-deletion, and black indicates no change (gray indicates poorly measured data). (*b*) Enlarged view of DNA copy number profiles across the X chromosome, shown for cell lines containing different numbers of X chromosomes.

identifying the starting position of the best and longest match of any DNA sequence represented in the corresponding UniGene cluster (10) against the "Golden Path" genome assembly (http://genome.ucsc.edu/; Oct 7, 2000 Freeze). For UniGene clusters represented by multiple arrayed elements, mean fluorescence ratios (for all elements representing the same UniGene cluster) are reported. For mRNA measurements, fluorescence ratios are "mean-centered" (i.e., reported relative to the mean ratio across the 44 tumor samples). The data set described here can be accessed in its entirety in the supporting information.

## Results

We performed CGH on 44 predominantly locally advanced, primary breast tumors and 10 breast cancer cell lines, using cDNA microarrays containing 6,691 different mapped human genes (Fig. 1*a*; also see *Materials and Methods* for details of microarray hybridizations). To take full advantage of the improved spatial resolution of array CGH, we ordered (fluorescence ratios for) the 6,691 cDNAs according to the "Golden Path" (http://genome.ucsc.edu/) genome assembly of the draft human genome sequences (11). In so doing, arrayed cDNAs not only themselves represent genes of potential interest (e.g., candidate oncogenes within amplicons), but also provide precise genetic landmarks for chromosomal regions of amplification and deletion. Parallel analysis of DNA from cell lines containing different numbers of X chromosomes (Fig. 1*b*), as we did before (7), demonstrated the sensitivity of our method to detect single-copy loss (45, XO), and 1.5- (47,XXX), 2- (48,XXXX), or 2.5-fold (49,XXXXX) gains (also see Fig. 5, which is published as supporting information on the PNAS web site). Fluorescence ratios were linearly proportional to copy number ratios, which were slightly underestimated, in agreement with previous observations (7). Numerous DNA copy number alterations were evident in both the breast cancer cell lines and primary tumors (Fig. 1*a*), detected in the tumors despite the presence of euploid non-tumor cell types; the magnitudes of the observed changes were generally lower in the tumor samples. DNA copy-number alterations were found in every cancer cell line and tumor, and on every human chromosome in at least one sample. Recurrent regions of DNA copy number gain and loss were readily identifiable. For example, gains within 1q, 8q, 17q, and 20q were observed in a high proportion of breast cancer cell lines/tumors (90%/69%, 100%/47%, 100%/60%, and 90%/44%, respectively), as were losses within 1p, 3p, 8p, and 13q (80%/24%, 80%/22%, 80%/22%, and 70%/18%, respectively), consistent with published cytogenetic studies (refs. 2–4; a complete listing of gains/losses is provided in Tables 2 and 3, which are published as supporting information on the PNAS web site). The total

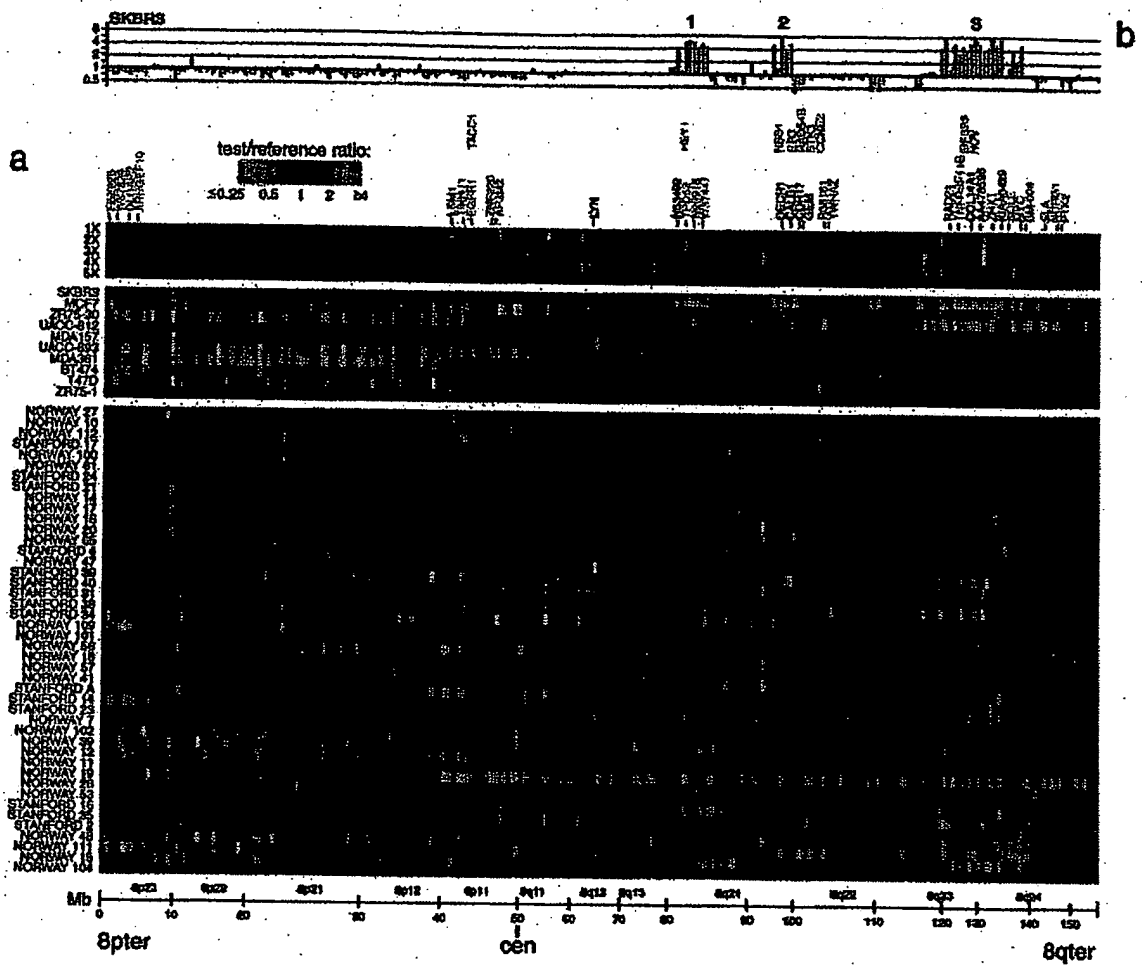



Fig. 2. DNA copy number alteration across chromosome 8 by array CGH. (*a*) DNA copy number profiles are illustrated for cell lines containing different numbers of X chromosomes, for breast cancer cell lines, and for breast tumors. Breast cancer cell lines and tumors are separately ordered by hierarchical clustering to highlight recurrent copy number changes. The 241 genes present on the microarrays and mapping to chromosome 8 are ordered by position along the chromosome. Fluorescence ratios (test/reference) are depicted by a $\log_2$ pseudocolor scale (indicated). Selected genes are indicated with color-coded text (red, increased; green, decreased; black, no change; gray, not well measured) to reflect correspondingly altered mRNA levels (observed in the majority of the subset of samples displaying the DNA copy number change). The map positions for genes of interest that are not represented on the microarray are indicated in the row above those genes represented on the array. (*b*) Graphical display of DNA copy number profile for breast cancer cell line SKBR3. Fluorescence ratios (tumor/normal) are plotted on a $\log_2$ scale for chromosome 8 genes, ordered along the chromosome.

number of genomic alterations (gains and losses) was found to be significantly higher in breast tumors that were high grade ($P = 0.008$), consistent with published CGH data (3), estrogen receptor negative ($P = 0.04$), and harboring TP53 mutations ($P = 0.0006$) (see Table 4, which is published as supporting information on the PNAS web site).

The improved spatial resolution of our array CGH analysis is illustrated for chromosome 8, which displayed extensive DNA copy number alteration in our series. A detailed view of the variation in the copy number of 241 genes mapping to chromosome 8 revealed multiple regions of recurrent amplification; each of these potentially harbors a different known or previously uncharacterized oncogene (Fig. 2*a*). The complexity of amplicon structure is most easily appreciated in the breast cancer cell line SKBR3. Although a conventional CGH analysis of 8q in SKBR3 identified only two distinct regions of amplification (12), we observed three distinct regions of high-level amplification (labeled 1–3 in Fig. 2*b*). For each of these regions we can define the boundaries of the interval recurrently amplified in the tumors we examined; in each case, known or plausible candidate oncogenes can be identified (a description of these regions, as well as the recurrently amplified regions on chromosomes 17 and 20, can be found in Figs. 6 and 7, which are published as supporting information on the PNAS web site).

For a subset of breast cancer cell lines and tumors (4 and 37, respectively), and a subset of arrayed genes (6,095), mRNA levels were quantitatively measured in parallel by using cDNA microarrays (8). The parallel assessment of mRNA levels is useful in the interpretation of DNA copy number changes. For example, the highly amplified genes that are also highly expressed are the strongest candidate oncogenes within an amplicon. Perhaps more significantly, our parallel analysis of DNA copy number changes and mRNA levels provides us the opportunity to assess the global impact of widespread DNA copy number alteration on gene expression in tumor cells.

A strong influence of DNA copy number on gene expression is evident in an examination of the pseudocolor representations

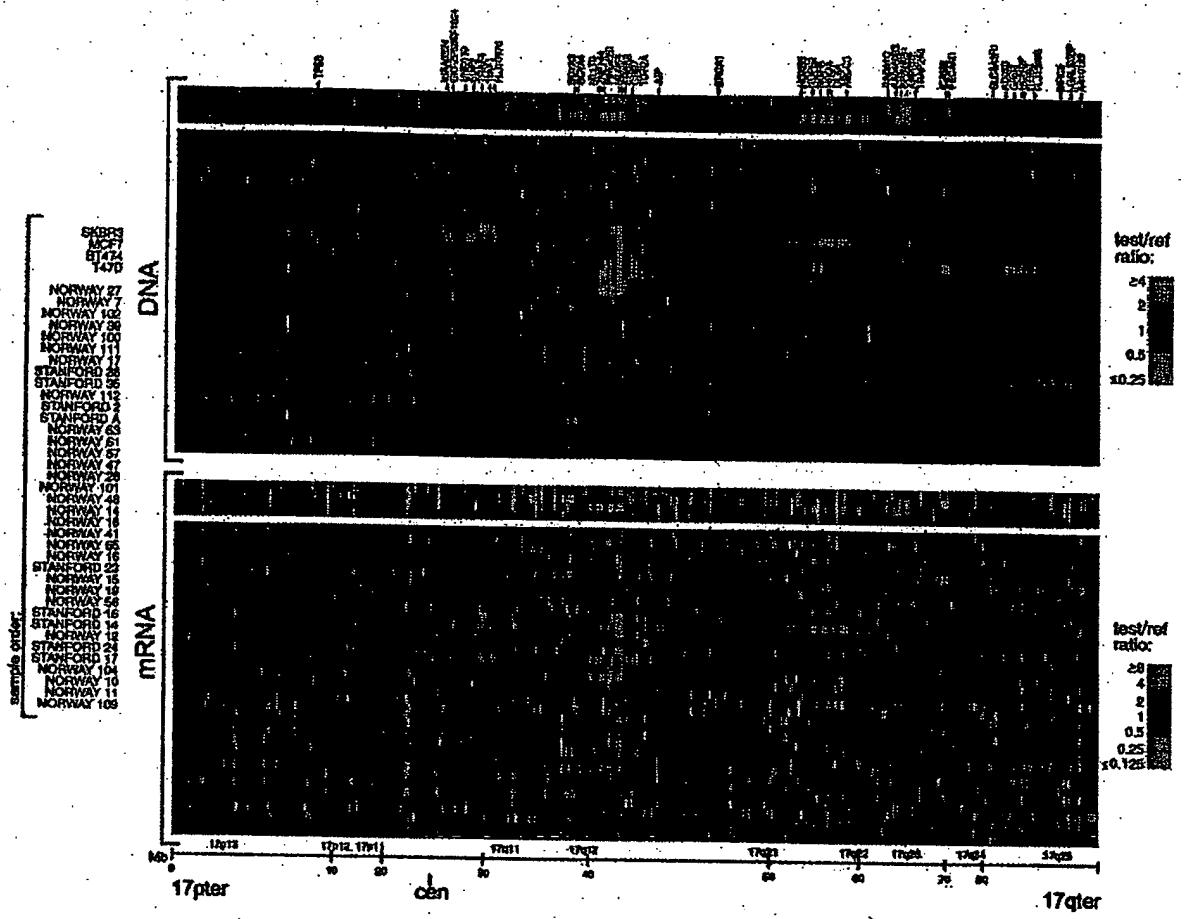



Fig. 3. Concordance between DNA copy number and gene expression across chromosome 17. DNA copy number alteration (*Upper*) and mRNA levels (*Lower*) are illustrated for breast cancer cell lines and tumors. Breast cancer cell lines and tumors are separately ordered by hierarchical clustering (*Upper*), and the identical sample order is maintained (*Lower*). The 354 genes present on the microarrays and mapping to chromosome 17, and for which both DNA copy number and mRNA levels were determined, are ordered by position along the chromosome; selected genes are indicated in color-coded text (see Fig. 2 legend). Fluorescence ratios (test/reference) are depicted by separate $\log_2$ pseudocolor scales (indicated).

of DNA copy number and mRNA levels for genes on chromosome 17 (Fig. 3). The overall patterns of gene amplification and elevated gene expression are quite concordant; i.e., a significant fraction of highly amplified genes appear to be correspondingly highly expressed. The concordance between high-level amplification and increased gene expression is not restricted to chromosome 17. Genome-wide, of 117 high-level DNA amplifications (fluorescence ratios >4, and representing 91 different genes), 62% (representing 54 different genes; see Table 5, which is published as supporting information on the PNAS web site) are found associated with at least moderately elevated mRNA levels (mean-centered fluorescence ratios >2), and 42% (representing 36 different genes) are found associated with comparably highly elevated mRNA levels (mean-centered fluorescence ratios >4).

To determine the extent to which DNA deletion and lower-level amplification (in addition to high-level amplification) are also associated with corresponding alterations in mRNA levels, we performed three separate analyses on the complete data set (4 cell lines and 37 tumors, across 6,095 genes). First, we determined the average mRNA levels for each of five classes of genes, representing DNA deletion, no change, and low-, medium-, and high-level amplification (Fig. 4*a*). For both the breast cancer cell lines and tumors, average mRNA levels tracked with DNA copy number across all five classes, in a statistically significant fashion ($P$ values for pair-wise Student's $t$ tests comparing adjacent classes: cell lines, $4 \times 10^{-49}$, $1 \times 10^{-49}$, $5 \times 10^{-5}$, $1 \times 10^{-2}$; tumors, $1 \times 10^{-43}$, $1 \times 10^{-214}$, $5 \times 10^{-41}$, $1 \times 10^{-4}$). A linear regression of the average log(DNA copy number), for each class, against average log(mRNA level) demonstrated that on average, a 2-fold change in DNA copy number was accompanied by 1.4- and 1.5-fold changes in mRNA level for the breast cancer cell lines and tumors, respectively (Fig. 4*a*, regression line not shown). Second, we characterized the distribution of the 6,095 correlations between DNA copy number and mRNA level, each across the 37 tumor samples (Fig. 4*b*). The distribution of correlations forms a normal-shaped curve, but with the peak markedly shifted in the positive direction from zero. This shift is statistically significant, as evidenced in a plot of observed vs. expected correlations (Fig. 4*c*), and reflects a pervasive global influence of DNA copy number alterations on gene expression. Notably, the highest correlations between DNA copy number and mRNA level (the right tail of the distribution in Fig. 4*b*) comprise both amplified and deleted genes (data not shown). Third, we used a linear regression model to estimate the fraction of all variation measured in mRNA levels among the 37

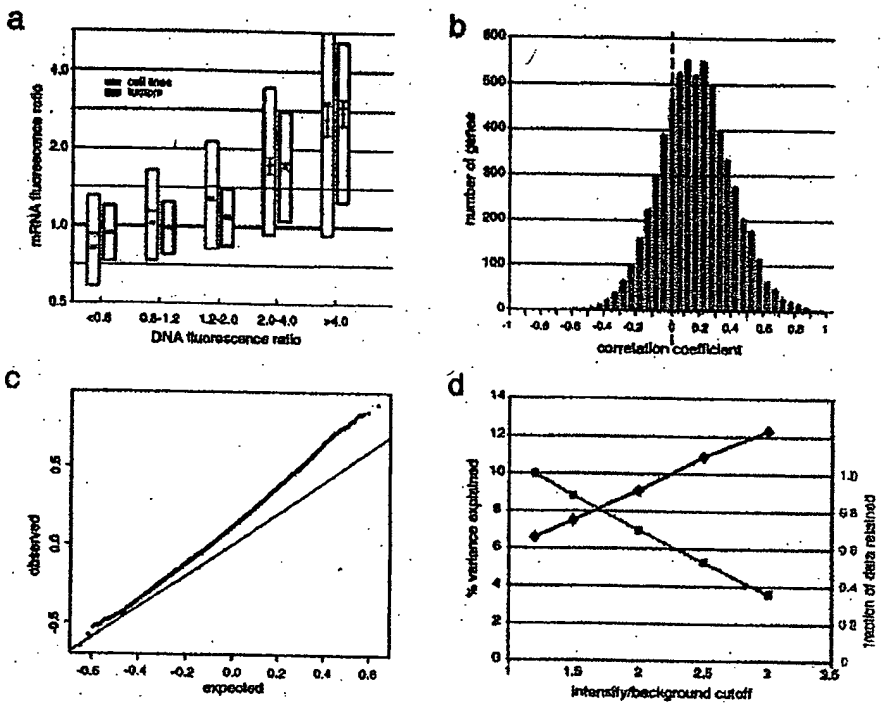



Fig. 4. Genome-wide influence of DNA copy number alterations on mRNA levels. (*a*) For breast cancer cell lines (gray) and tumor samples (black), both mean-centered mRNA fluorescence ratio ($log_2$ scale) quartiles (box plots indicate 25th, 50th, and 75th percentile) and averages (diamonds; *Y*-value error bars indicate standard errors of the mean) are plotted for each of five classes of genes, representing DNA deletion (tumor/normal ratio < 0.8), no change (0.8–1.2), low- (1.2–2), medium- (2–4), and high-level (>4) amplification. *P* values for pair-wise Student's *t* tests, comparing averages between adjacent classes (moving left to right), are $4 \times 10^{-49}$, $1 \times 10^{-49}$, $5 \times 10^{-5}$, $1 \times 10^{-2}$ (cell lines), and $1 \times 10^{-43}$, $1 \times 10^{-214}$, $5 \times 10^{-41}$, $1 \times 10^{-4}$ (tumors). (*b*) Distribution of correlations between DNA copy number and mRNA levels, for 6,095 different human genes across 37 breast tumor samples. (*c*) Plot of observed versus expected correlation coefficients. The expected values were obtained by randomization of the sample labels in the DNA copy number data set. The line of unity is indicated. (*d*) Percent variance in gene expression (among tumors) directly explained by variation in gene copy number. Percent variance explained (black line) and fraction of data retained (gray line) are plotted for different fluorescence intensity/background (a rough surrogate for signal/noise) cutoff values. Fraction of data retained is relative to the 1.2 intensity/background cutoff. Details of the linear regression model used to estimate the fraction of variation in gene expression attributable to underlying DNA copy number alteration can be found in the supporting information (see *Estimating the Fraction of Variation in Gene Expression Attributable to Underlying DNA Copy Number Alteration*).

tumors that could be attributed to underlying variation in DNA copy number. From this analysis, we estimate that, overall, about 7% of all of the observed variation in mRNA levels can be explained directly by variation in copy number of the altered genes (Fig. 4*d*). We can reduce the effects of experimental measurement error on this estimate by using only that fraction of the data most reliably measured (fluorescence intensity/background >3); using that data, our estimate of the percent variation in mRNA levels directly attributed to variation in gene copy number increases to 12% (Fig. 4*d*). This still undoubtedly represents a significant underestimate, as the observed variation in global gene expression is affected not only by true variation in the expression programs of the tumor cells themselves, but also by the variable presence of non-tumor cell types within clinical samples.

## Discussion

This genome-wide, array CGH analysis of DNA copy number alteration in a series of human breast tumors demonstrates the usefulness of defining amplicon boundaries at high resolution (gene-by-gene), and quantitatively measuring amplicon shape, to assist in locating and identifying candidate oncogenes. By analyzing mRNA levels in parallel, we have also discovered that changes in DNA copy number have a large, pervasive, direct effect on global gene expression patterns in both breast cancer cell lines and tumors. Although the DNA microarrays used in our analysis may display a bias toward characterized and/or highly expressed genes, because we are examining such a large fraction of the genome (approximately 20% of all human genes), and because, as detailed above, we are likely underestimating the contribution of DNA copy number changes to altered gene expression, we believe our findings are likely to be generalizable (but would nevertheless still be remarkable if only applicable to this set of ~6,100 genes).

In budding yeast, aneuploidy has been shown to result in chromosome-wide gene expression biases (13). Two recent studies have begun to examine the global relationship between DNA copy number and gene expression in cancer cells. In agreement with our findings, Phillips *et al.* (14) have shown that with the acquisition of tumorigenicity in an immortalized prostate epithelial cell line, new chromosomal gains and losses resulted in a statistically significant respective increase and decrease in the average expression level of involved genes. In contrast, Platzer *et al.* (15) recently reported that in metastatic colon tumors only ~4% of genes within amplified regions were found more highly (>2-fold) expressed, when compared with normal colonic epithelium. This report differs substantially from our finding that 62% of highly amplified genes in breast cancer exhibit at least 2-fold increased expression. These contrasting findings may reflect methodological differences between the

studies. For example, the study of Platzer *et al.* (15) may have systematically under-measured gene expression changes. In this regard it is remarkable that only 14 transcripts of many thousand residing within unamplified chromosomal regions were found to exhibit at least 4-fold altered expression in metastatic colon cancer. Additionally, their reliance on lower-resolution chromosomal CGH may have resulted in poorly delimiting the boundaries of high-complexity amplicons, effectively overcalling regions with amplification. Alternatively, the contrasting findings for amplified genes may represent real biological differences between breast and metastatic colon tumors; resolution of this issue will require further studies.

Our finding that widespread DNA copy number alteration has a large, pervasive and direct effect on global gene expression patterns in breast cancer has several important implications. First, this finding supports a high degree of copy number-dependent gene expression in tumors. Second, it suggests that most genes are not subject to specific autoregulation or dosage compensation. Third, this finding cautions that elevated expression of an amplified gene cannot alone be considered strong independent evidence of a candidate oncogene's role in tumorigenesis. In our study, fully 62% of highly amplified genes demonstrated moderately or highly elevated expression. This highlights the importance of high-resolution mapping of amplicon boundaries and shape [to identify the "driving" gene(s) within amplicons (16)], on a large number of samples, in addition to functional studies. Fourth, this finding suggests that analyzing the genomic distribution of expressed genes, even within existing microarray gene expression data sets, may permit the inference of DNA copy number aberration, particularly aneuploidy (where gene expression can be averaged across large chromosomal regions; see Fig. 3 and supporting information). Fifth, this finding implies that a substantial portion of the phenotypic uniqueness (and by extension, the heterogeneity in clinical behavior) among patients' tumors may be traceable to underlying variation in DNA copy number. Sixth, this finding supports a possible role for widespread DNA copy number alteration in tumorigenesis (17, 18), beyond the amplification of specific oncogenes and deletion of specific tumor suppressor genes. Widespread DNA copy number alteration, and the concomitant widespread imbalance in gene expression, might disrupt critical stochiometric relationships in cell metabolism and physiology (e.g., proteosome, mitotic spindle), possibly promoting further chromosomal instability and directly contributing to tumor development or progression. Finally, our findings suggest the possibility of cancer therapies that exploit specific or global imbalances in gene expression in cancer.

We thank the many members of the P.O.B. and D.B. labs for helpful discussions. J.R.P. was a Howard Hughes Medical Institute Physician Postdoctoral Fellow during a portion of this work. P.O.B. is a Howard Hughes Medical Institute Associate Investigator. This work was supported by grants from the National Institutes of Health, the Howard Hughes Medical Institute, the Norwegian Cancer Society, and the Norwegian Research Council.

1. Kallioniemi, A., Kallioniemi, O. P., Sudar, D., Rutovitz, D., Gray, J. W., Waldman, F. & Pinkel, D. (1992) *Science* 258, 818–821.
2. Kallioniemi, A., Kallioniemi, O. P., Piper, J., Tanner, M., Stokke, T., Chen, L., Smith, H. S., Pinkel, D., Gray, J. W. & Waldman, F. M. (1994) *Proc. Natl. Acad. Sci. USA* 91, 2156–2160.
3. Tirkkonen, M., Tanner, M., Karhu, R., Kallioniemi, A., Isola, J. & Kallioniemi, O. P. (1998) *Genes Chromosomes Cancer* 21, 177–184.
4. Forozan, F., Mahlamaki, E. H., Monni, O., Chen, Y., Veldman, R., Jiang, Y., Gooden, G. C., Ethier, S. P., Kallioniemi, A. & Kallioniemi, O. P. (2000) *Cancer Res.* 60, 4519–4525.
5. Solinas-Toldo, S., Lampel, S., Stilgenbauer, S., Nicktolenko, J., Benner, A., Dohner, H., Cremer, T. & Lichter, P. (1997) *Genes Chromosomes Cancer* 20, 399–407.
6. Pinkel, D., Segraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W. L., Chen, C., Zhai, Y., *et al.* (1998) *Nat. Genet.* 20, 207–211.
7. Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D. & Brown, P. O. (1999) *Nat. Genet.* 23, 41–46.
8. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., *et al.* (2000) *Nature (London)* 406, 747–752.
9. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* 98, 10869–10874.
10. Schuler, G. D. (1997) *J. Mol. Med.* 75, 694–698.
11. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001) *Nature (London)* 409, 860–921.
12. Fejzo, M. S., Godfrey, T., Chen, C., Waldman, F. & Gray, J. W. (1998) *Genes Chromosomes Cancer* 23, 105–113.
13. Hughes, T. R., Roberts, C. J., Dai, H., Jones, A. R., Meyer, M. R., Slade, D., Burchard, J., Dow, S., Ward, T. R., Kidd, M. J., *et al.* (2000) *Nat. Genet.* 25, 333–337.
14. Phillips, J. L., Hayward, S. W., Wang, Y., Vasselli, J., Pavlovich, C., Padilla-Nash, H., Pezullo, J. R., Ghadimi, B. M., Grossfeld, G. D., Rivera, A., *et al.* (2001) *Cancer Res.* 61, 8143–8149.
15. Platzer, P., Upender, M. B., Wilson, K., Willis, J., Lutterbaugh, J., Nosrati, A., Willson, J. K., Mack, D., Ried, T. & Markowitz, S. (2002) *Cancer Res.* 62, 1134–1138.
16. Albertson, D. G., Ylstra, B., Segraves, R., Collins, C., Dairkee, S. H., Kowbel, D., Kuo, W. L., Gray, J. W. & Pinkel, D. (2000) *Nat. Genet.* 25, 144–146.
17. Li, R., Yerganian, G., Duesberg, P., Kraemer, A., Willer, A., Rausch, C. & Hehlmann, R. (1997) *Proc. Natl. Acad. Sci. USA* 94, 14506–14511.
18. Rasnick, D. & Duesberg, P. H. (1999) *Biochem. J.* 340, 621–630.

# MOLECULAR BIOLOGY OF THE CELL

fourth edition

Bruce Alberts

Alexander Johnson

Julian Lewis

Martin Raff

Keith Roberts

Peter Walter

GS Garland Science
Taylor & Francis Group

*Garland*
Vice President: Denise Schanck
Managing Editor: Sarah Gibbs
Senior Editorial Assistant: Kirsten Jenner
Managing Production Editor: Emma Hunt
Proofreader and Layout: Emma Hunt
Production Assistant: Angela Bennett
Text Editors: Marjorie Singer Anderson and Betsy Dilernia
Copy Editor: Bruce Goatly
Word Processors: Fran Dependahl, Misty Landers and Carol Winter
Designer: Blink Studio, London
Illustrator: Nigel Orme
Indexer: Janine Ross and Sherry Granum
Manufacturing: Nigel Eyre and Marion Morrow

**Bruce Alberts** received his Ph.D. from Harvard University and is President of the National Academy of Sciences and Professor of Biochemistry and Biophysics at the University of California, San Francisco. **Alexander Johnson** received his Ph.D. from Harvard University and is a Professor of Microbiology and Immunology at the University of California, San Francisco. **Julian Lewis** received his D.Phil. from the University of Oxford and is a Principal Scientist at the Imperial Cancer Research Fund, London. **Martin Raff** received his M.D. from McGill University and is at the Medical Research Council Laboratory for Molecular Cell Biology and Cell Biology Unit and in the Biology Department at University College London. **Keith Roberts** received his Ph.D. from the University of Cambridge and is Associate Research Director at the John Innes Centre, Norwich. **Peter Walter** received his Ph.D. from The Rockefeller University in New York and is Professor and Chairman of the Department of Biochemistry and Biophysics at the University of California, San Francisco, and an Investigator of the Howard Hughes Medical Institute.



**Library of Congress Cataloging-in-Publicaton Data**
Molecular biology of the cell / Bruce Alberts ... [et al.].-- 4th ed.
p. cm
Includes bibliographical references and index.
ISBN 0-8153-3218-1 (hardbound) -- ISBN 0-8153-4072-9 (pbk.)
1. Cytology. 2. Molecular biology. I. Alberts, Bruce.
[DNLM: 1. Cells. 2. Molecular Biology. ]
QH581.2 .M64 2002
571.6--dc21
2001054471 CIP

Published by Garland Science, a member of the Taylor & Francis Group, 29 West 35th Street, New York, NY 10001-2299

Printed in the United States of America

15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

*Cell Biology Interactive*
Artistic and Scientific Direction: Peter Walter
Narrated by: Julie Theriot
Production, Design, and Development: Mike Morales

**Front cover** Human Genome: Reprinted by permission from *Nature,* International Human Genome Sequencing Consortium, 409:860–921, 2001 © Macmillan Magazines Ltd. Adapted from an image by Francis Collins, NHGRI; Jim Kent, UCSC; Ewan Birney, EBI; and Darryl Leja, NHGRI; showing a portion of Chromosome 1 from the initial sequencing of the human genome.

**Back cover** In 1967, the British artist Peter Blake created a design classic. Nearly 35 years later Nigel Orme (illustrator), Richard Denyer (photographer), and the authors have together produced an affectionate tribute to Mr Blake's image. With its gallery of icons and influences, its assembly created almost as much complexity, intrigue and mystery as the original. *Drosophila, Arabidopsis,* Dolly and the assembled company tempt you to dip inside where, as in the original, "a splendid time is guaranteed for all." (Gunter Blobel, courtesy of The Rockefeller University; Marie Curie, Keystone Press Agency Inc; Darwin bust, by permission of the President and Council of the Royal Society; Rosalind Franklin, courtesy of Cold Spring Harbor Laboratory Archives; Dorothy Hodgkin, © The Nobel Foundation, 1964; James Joyce, etching by Peter Blake; Robert Johnson, photo booth self-portrait early 1930s, © 1986 Delta Haze Corporation all rights reserved, used by permission; Albert L. Lehninger, (unidentified photographer) courtesy of The Alan Mason Chesney Medical Archives of The Johns Hopkins Medical Institutions; Linus Pauling, from Ava Helen and Linus Pauling Papers, Special Collections, Oregon State University; Nicholas Poussin, courtesy of ArtToday.com; Barbara McClintock, © David Micklos, 1983; Andrei Sakharov, courtesy of Elena Bonner; Frederick Sanger, © The Nobel Foundation, 1958.)

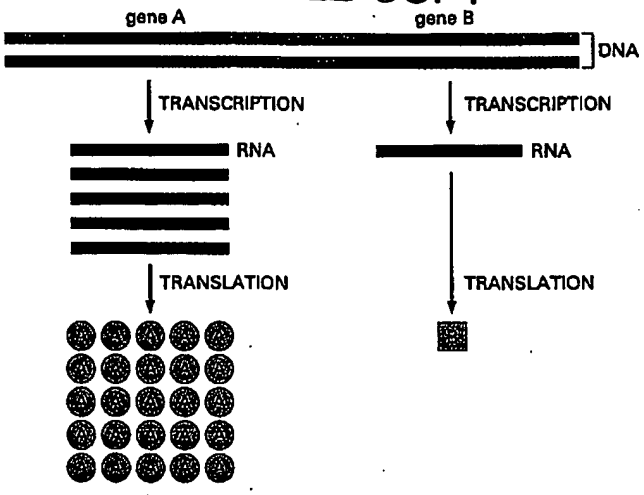



**Figure 6–3 Genes can be expressed with different efficiencies.** Gene A is transcribed and translated much more efficiently than gene B. This allows the amount of protein A in the cell to be much greater than that of protein B.

## FROM DNA TO RNA

Transcription and translation are the means by which cells read out, or express, the genetic instructions in their genes. Because many identical RNA copies can be made from the same gene, and each RNA molecule can direct the synthesis of many identical protein molecules, cells can synthesize a large amount of protein rapidly when necessary. But each gene can also be transcribed and translated with a different efficiency, allowing the cell to make vast quantities of some proteins and tiny quantities of others (Figure 6–3). Moreover, as we see in the next chapter, a cell can change (or regulate) the expression of each of its genes according to the needs of the moment—most obviously by controlling the production of its RNA.

### Portions of DNA Sequence Are Transcribed into RNA

The first step a cell takes in reading out a needed part of its genetic instructions is to copy a particular portion of its DNA nucleotide sequence—a gene—into an RNA nucleotide sequence. The information in RNA, although copied into another chemical form, is still written in essentially the same language as it is in DNA—the language of a nucleotide sequence. Hence the name **transcription.**

Like DNA, RNA is a linear polymer made of four different types of nucleotide subunits linked together by phosphodiester bonds (Figure 6–4). It differs from DNA chemically in two respects: (1) the nucleotides in RNA are *ribonucleotides*—that is, they contain the sugar ribose (hence the name *ribo*nucleic acid) rather than deoxyribose; (2) although, like DNA, RNA contains the bases adenine (A), guanine (G), and cytosine (C), it contains the base uracil (U) instead of the thymine (T) in DNA. Since U, like T, can base-pair by hydrogen-bonding with A (Figure 6–5), the complementary base-pairing properties described for DNA in Chapters 4 and 5 apply also to RNA (in RNA, G pairs with C, and A pairs with U). It is not uncommon, however, to find other types of base pairs in RNA: for example, G pairing with U occasionally.

Despite these small chemical differences, DNA and RNA differ quite dramatically in overall structure. Whereas DNA always occurs in cells as a double-stranded helix, RNA is single-stranded. RNA chains therefore fold up into a variety of shapes, just as a polypeptide chain folds up to form the final shape of a protein (Figure 6–6). As we see later in this chapter, the ability to fold into complex three-dimensional shapes allows some RNA molecules to have structural and catalytic functions.

### Transcription Produces RNA Complementary to One Strand of DNA

All of the RNA in a cell is made by DNA transcription, a process that has certain similarities to the process of DNA replication discussed in Chapter 5.

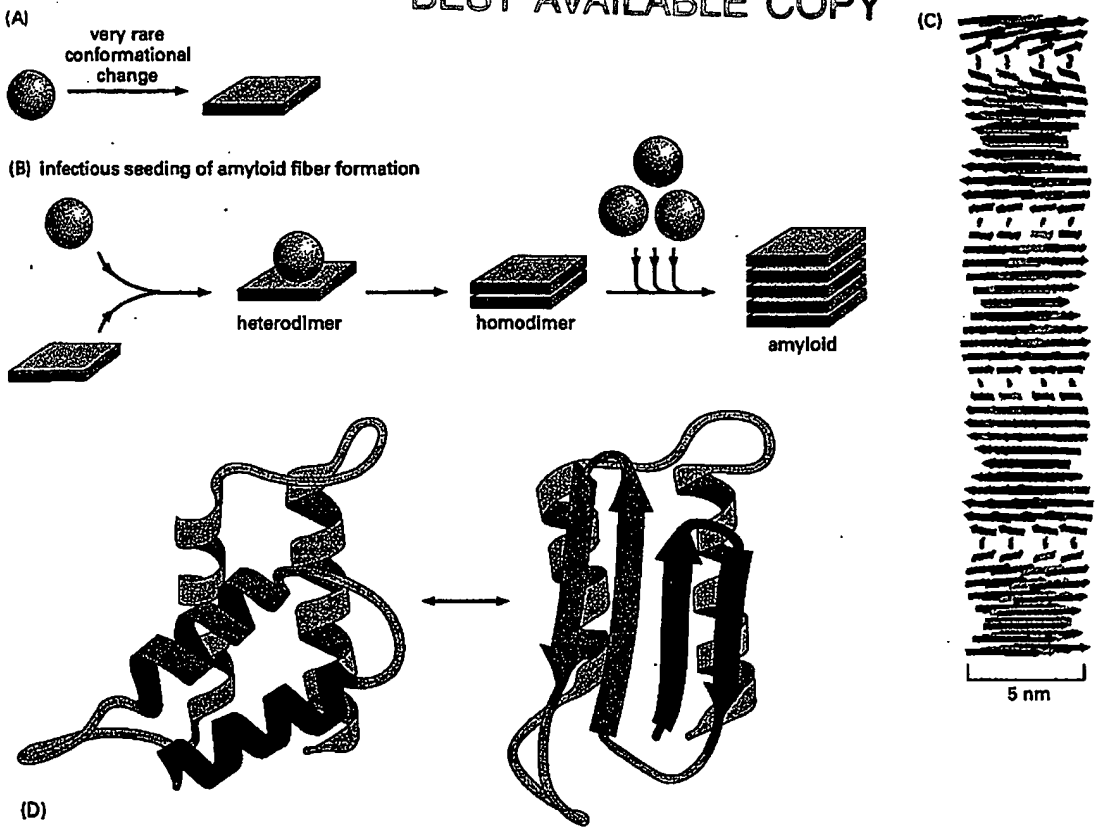



**Figure 6–89 Protein aggregates that cause human disease.** (A) Schematic illustration of the type of conformational change in a protein that produces material for a cross-beta filament. (B) Diagram illustrating the self-infectious nature of the protein aggregation that is central to prion diseases. PrP is highly unusual because the misfolded version of the protein, called PrP*, induces the normal PrP protein it contacts to change its conformation, as shown. Most of the human diseases caused by protein aggregation are caused by the overproduction of a variant protein that is especially prone to aggregation, but because this structure is not infectious in this way, it cannot spread from one animal to another. (C) Drawing of a cross-beta filament, a common type of protease-resistant protein aggregate found in a variety of human neurological diseases. Because the hydrogen-bond interactions in a β sheet form between polypeptide backbone atoms (see Figure 3–9), a number of different abnormally folded proteins can produce this structure. (D) One of several possible models for the conversion of PrP to PrP*, showing the likely change of two α-helices into four β-strands. Although the structure of the normal protein has been determined accurately, the structure of the infectious form is not yet known with certainty because the aggregation has prevented the use of standard structural techniques. (C, courtesy of Louise Serpell, adapted from M. Sunde et al., *J. Mol. Biol.* 273:729–739, 1997; D, adapted from S.B. Prusiner, *Trends Biochem. Sci.* 21:482–487, 1996.)

animals and humans. It can be dangerous to eat the tissues of animals that contain PrP*, as witnessed most recently by the spread of BSE (commonly referred to as the "mad cow disease") from cattle to humans in Great Britain.

Fortunately, in the absence of PrP*, PrP is extraordinarily difficult to convert to its abnormal form. Although very few proteins have the potential to misfold into an infectious conformation, a similar transformation has been discovered to be the cause of an otherwise mysterious "protein-only inheritance" observed in yeast cells.

## There Are Many Steps From DNA to Protein

We have seen so far in this chapter that many different types of chemical reactions are required to produce a properly folded protein from the information contained in a gene (Figure 6–90). The final level of a properly folded protein in a cell therefore depends upon the efficiency with which each of the many steps is performed.

We discuss in Chapter 7 that cells have the ability to change the levels of their proteins according to their needs. In principle, any or all of the steps in Fig-

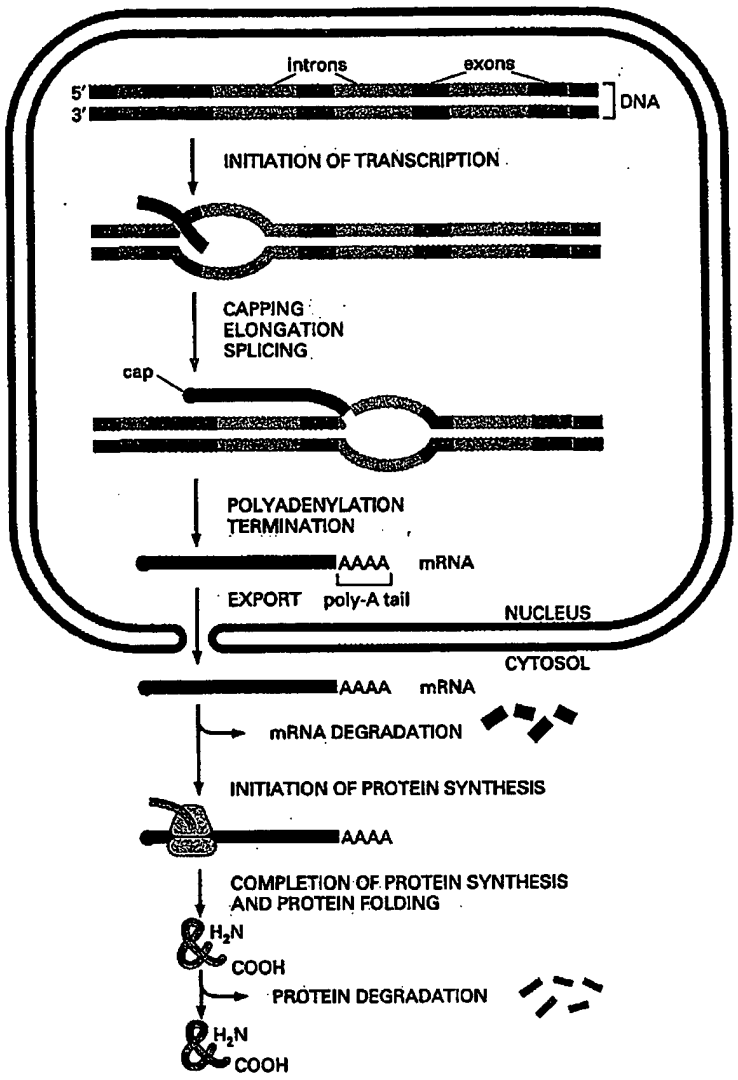



**Figure 6–90 The production of a protein by a eucaryotic cell.** The final level of each protein in a eucaryotic cell depends upon the efficiency of each step depicted.

ure 6–90) could be regulated by the cell for each individual protein. However, as we shall see in Chapter 7, the initiation of transcription is the most common point for a cell to regulate the expression of each of its genes. This makes sense, inasmuch as the most efficient way to keep a gene from being expressed is to block the very first step—the transcription of its DNA sequence into an RNA molecule.

## Summary

***The translation of the nucleotide sequence of an mRNA molecule into protein takes place in the cytoplasm on a large ribonucleoprotein assembly called a ribosome. The amino acids used for protein synthesis are first attached to a family of tRNA molecules, each of which recognizes, by complementary base-pair interactions, particular sets of three nucleotides in the mRNA (codons). The sequence of nucleotides in the mRNA is then read from one end to the other in sets of three according to the genetic code.***

***To initiate translation, a small ribosomal subunit binds to the mRNA molecule at a start codon (AUG) that is recognized by a unique initiator tRNA molecule. A large ribosomal subunit binds to complete the ribosome and begin the elongation phase of protein synthesis. During this phase, aminoacyl tRNAs—each bearing a specific amino acid bind sequentially to the appropriate codon in mRNA by forming complementary base pairs with the tRNA anticodon. Each amino acid is added to the C-terminal end of the growing polypeptide by means of a cycle of three sequential***

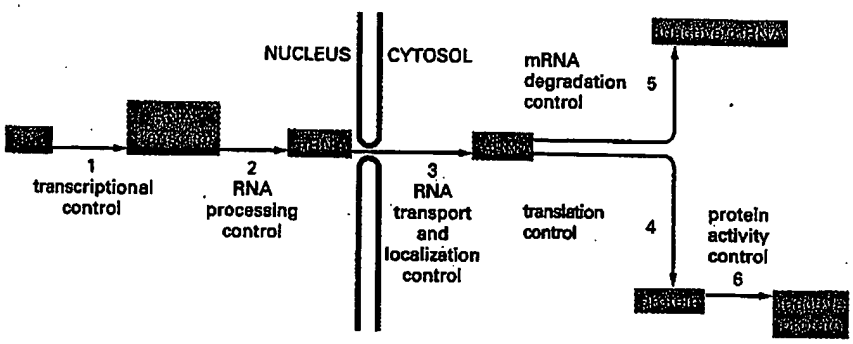



**Figure 7–5 Six steps at which eucaryotic gene expression can be controlled.** Controls that operate at steps 1 through 5 are discussed in this chapter. Step 6, the regulation of protein activity, includes reversible activation or inactivation by protein phosphorylation (discussed in Chapter 3) as well as irreversible inactivation by proteolytic degradation (discussed in Chapter 6).

## Gene Expression Can Be Regulated at Many of the Steps in the Pathway from DNA to RNA to Protein

If differences among the various cell types of an organism depend on the particular genes that the cells express, at what level is the control of gene expression exercised? As we saw in the last chapter, there are many steps in the pathway leading from DNA to protein, and all of them can in principle be regulated. Thus a cell can control the proteins it makes by (1) controlling when and how often a given gene is transcribed (**transcriptional control**), (2) controlling how the RNA transcript is spliced or otherwise processed (**RNA processing control**), (3) selecting which completed mRNAs in the cell nucleus are exported to the cytosol and determining where in the cytosol they are localized (**RNA transport and localization control**), (4) selecting which mRNAs in the cytoplasm are translated by ribosomes (**translational control**), (5) selectively destabilizing certain mRNA molecules in the cytoplasm (**mRNA degradation control**), or (6) selectively activating, inactivating, degrading, or compartmentalizing specific protein molecules after they have been made (**protein activity control**) (Figure 7–5).

For most genes transcriptional controls are paramount. This makes sense because, of all the possible control points illustrated in Figure 7–5, only transcriptional control ensures that the cell will not synthesize superfluous intermediates. In the following sections we discuss the DNA and protein components that perform this function by regulating the initiation of gene transcription. We shall return at the end of the chapter to the additional ways of regulating gene expression.

## Summary

***The genome of a cell contains in its DNA sequence the information to make many thousands of different protein and RNA molecules. A cell typically expresses only a fraction of its genes, and the different types of cells in multicellular organisms arise because different sets of genes are expressed. Moreover, cells can change the pattern of genes they express in response to changes in their environment, such as signals from other cells. Although all of the steps involved in expressing a gene can in principle be regulated, for most genes the initiation of RNA transcription is the most important point of control.***

# DNA-BINDING MOTIFS IN GENE REGULATORY PROTEINS

How does a cell determine which of its thousands of genes to transcribe? As mentioned briefly in Chapters 4 and 6, the transcription of each gene is controlled by a regulatory region of DNA relatively near the site where transcription begins. Some regulatory regions are simple and act as switches that are thrown by a single signal. Many others are complex and act as tiny microprocessors, responding to a variety of signals that they interpret and integrate to switch the neighboring gene on or off. Whether complex or simple, these switching devices

occur in the germ line, the cell lineage that gives rise to sperm or eggs. Most of the DNA in vertebrate germ cells is inactive and highly methylated. Over long periods of evolutionary time, the methylated CG sequences in these inactive regions have presumably been lost through spontaneous deamination events that were not properly repaired. However promoters of genes that remain active in the germ cell lineages (including most housekeeping genes) are kept unmethylated, and therefore spontaneous deaminations of Cs that occur within them can be accurately repaired. Such regions are preserved in modern day vertebrate cells as CG islands. In addition, any mutation of a CG sequence in the genome that destroyed the function or regulation of a gene in the adult would be selected against, and some CG islands are simply the result of a higher than normal density of critical CG sequences.

The mammalian genome contains an estimated 20,000 CG islands. Most of the islands mark the 5′ ends of transcription units and thus, presumably, of genes. The presence of CG islands often provides a convenient way of identifying genes in the DNA sequences of vertebrate genomes.

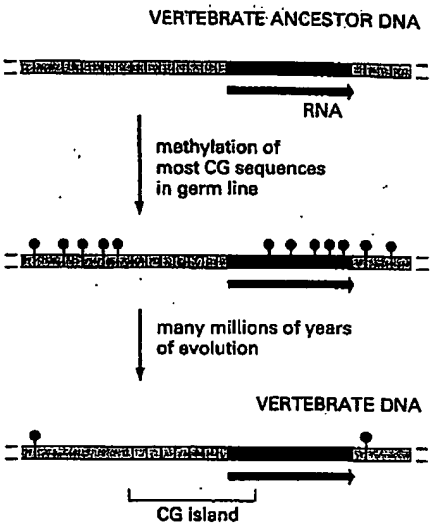



**Figure 7–86 A mechanism to explain both the marked overall deficiency of CG sequences and their clustering into CG islands in vertebrate genomes.** A *black line* marks the location of a CG dinucleotide in the DNA sequence, while a *red* "lollipop" indicates the presence of a methyl group on the CG dinucleotide. CG sequences that lie in regulatory sequences of genes that are transcribed in germ cells are unmethylated and therefore tend to be retained in evolution. Methylated CG sequences, on the other hand, tend to be lost through deamination of 5-methyl C to T, unless the CG sequence is critical for survival.

## Summary

***The many types of cells in animals and plants are created largely through mechanisms that cause different genes to be transcribed in different cells. Since many specialized animal cells can maintain their unique character through many cell division cycles and even when grown in culture, the gene regulatory mechanisms involved in creating them must be stable once established and heritable when the cell divides. These features endow the cell with a memory of its developmental history. Bacteria and yeasts provide unusually accessible model systems in which to study gene regulatory mechanisms. One such mechanism involves a competitive interaction between two gene regulatory proteins, each of which inhibits the synthesis of the other; this can create a flip-flop switch that switches a cell between two alternative patterns of gene expression. Direct or indirect positive feedback loops, which enable gene regulatory proteins to perpetuate their own synthesis, provide a general mechanism for cell memory. Negative feedback loops with programmed delays form the basis for cellular clocks.***

***In eucaryotes the transcription of a gene is generally controlled by combinations of gene regulatory proteins. It is thought that each type of cell in a higher eucaryotic organism contains a specific combination of gene regulatory proteins that ensures the expression of only those genes appropriate to that type of cell. A given gene regulatory protein may be active in a variety of circumstances and typically is involved in the regulation of many genes.***

***In addition to diffusible gene regulatory proteins, inherited states of chromatin condensation are also used by eucaryotic cells to regulate gene expression. An especially dramatic case is the inactivation of an entire X chromosome in female mammals. In vertebrates DNA methylation also functions in gene regulation, being used mainly as a device to reinforce decisions about gene expression that are made initially by other mechanisms. DNA methylation also underlies the phenomenon of genomic imprinting in mammals, in which the expression of a gene depends on whether it was inherited from the mother or the father.***

# POSTTRANSCRIPTIONAL CONTROLS

In principle, every step required for the process of gene expression could be controlled. Indeed, one can find examples of each type of regulation, although any one gene is likely to use only a few of them. Controls on the initiation of gene transcription are the predominant form of regulation for most genes. But other controls can act later in the pathway from DNA to protein to modulate the amount of gene product that is made. Although these **posttranscriptional controls**, which operate after RNA polymerase has bound to the gene's promoter and begun RNA synthesis, are less common than *transcriptional control*, for many genes they are crucial.

TECHNICAL UPDATE

FROM YOUR LABORATORY SERVICES PROVIDER

# HER-2/neu Breast Cancer Predictive Testing

*Julie Sanford Hanna, Ph.D. and Dan Mornin, M.D.*

EACH YEAR, OVER 182,000 WOMEN in the United States are diagnosed with breast cancer, and approximately 45,000 die of the disease.[1] Incidence appears to be increasing in the United States at a rate of roughly 2% per year. The reasons for the increase are unclear, but non-genetic risk factors appear to play a large role.[2]

Five-year survival rates range from approximately 65%-85%, depending on demographic group, with a significant percentage of women experiencing recurrence of their cancer within 10 years of diagnosis. One of the factors most predictive for recurrence once a diagnosis of breast cancer has been made is the number of axillary lymph nodes to which tumor has metastasized. Most node-positive women are given adjuvant therapy, which increases their survival. However, 20%-30% of patients without axillary node involvement also develop recurrent disease, and the difficulty lies in how to identify this high-risk subset of patients. These patients could benefit from increased surveillance, early intervention, and treatment.

Prognostic markers currently used in breast cancer recurrence prediction include tumor size, histological grade, steroid hormone receptor status, DNA ploidy, proliferative index, and cathepsin D status. Expression of growth factor receptors and over-expression of the HER-2/neu oncogene have also been identified as having value regarding treatment regimen and prognosis.

HER-2/neu (also known as c-erbB2) is an oncogene that encodes a transmembrane glycoprotein that is homologous to, but distinct from, the epidermal growth factor receptor. Numerous studies have indicated that high levels of expression of this protein are associated with rapid tumor growth, certain forms of therapy resistance, and shorter disease-free survival. The gene has been shown to be amplified and/or overexpressed in 10%-30% of invasive breast cancers and in 40%-60% of intraductal breast carcinoma.[3]

There are two distinct FDA-approved methods by which HER-2/neu status can be evaluated: immunohistochemistry (IHC, HercepTest™) and FISH (fluorescent in situ hybridization, PathVysion™ Kit). Both methods can be performed on archived and current specimens. The first method allows visual assessment of the amount of HER-2/neu protein present on the cell membrane. The latter method allows direct quantification of the level of gene amplification present in the tumor, enabling differentiation between low- versus high-amplification. At least one study has demonstrated a difference in recurrence risk in women younger than 40 years of age for low- versus high-amplified tumors (54.5% compared to 85.7%); this is compared to a recurrence rate of 16.7% for patients with no HER-2/neu gene amplification.[4] HER-2/neu status may be particularly important to establish in women with small (≤1 cm) tumor size.

The choice of methodology for determination of HER-2/neu status depends in part on the clinical setting. FDA approval for the Vysis FISH test was granted based on clinical trials involving 1549 node-positive patients. Patients received one of three different treatments consisting of different doses of cyclophosphamide, Adriamycin, and 5-fluorouracil (CAF). The study showed that patients with amplified HER-2/neu benefited from treatment with higher doses of adriamycin-based therapy, while those with normal HER-2/neu levels did not. The study therefore identified a sub-set of women, who because they did not benefit from more aggressive treatment, did not need to be exposed to the associated side effects. In addition, other evidence indicates that HER-2/neu amplification in node-negative patients can be used as an independent prognostic indicator for early recurrence, recurrent disease at any time and disease-related death.[5] Demonstration of HER-2/neu gene amplification by FISH has also been shown to be of value in predicting response to chemotherapy in stage-2 breast cancer patients.

Selection of patients for Herceptin© (Trastuzumab) monoclonal antibody therapy, however, is based upon demonstration of HER-2/neu protein overexpression using HercepTest™. Studies using Herceptin© in patients with metastatic breast cancer show an increase in time to disease progression, increased response rate to chemotherapeutic agents and a small increase in overall survival rate. The FISH assays have not yet been approved for this purpose, and studies looking at response to Herceptin© in patients with or without gene amplification status determined by FISH are in progress.

In general, FISH and IHC results correlate well. However, subsets of tumors are found which show discordant results; i.e., protein overexpression without gene amplification or lack of protein overexpression with gene amplification. The clinical significance of such results is unclear. Based on the above considerations, HER-2/neu testing at SHMC/PAML will utilize immunohistochemistry (HercepTest©) as a screen, followed by FISH in IHC-negative cases. Alternatively, either method may be ordered individually depending on the clinical setting or clinician preference.

## CPT code information

**HER-2/neu via IHC**

88342 (including interpretive report)

**HER-2/neu via FISH**

88271×2 Molecular cytogenetics, DNA probe, each
88274 Molecular cytogenetics, interphase in situ hybridization, analyze 25-99 cells
88291 Cytogenetics and molecular cytogenetics, interpretation and report

## Procedural Information

Immunohistochemistry is performed using the FDA-approved DAKO antibody kit, Herceptest©. The DAKO kit contains reagents required to complete a two-step immunohistochemical staining procedure for routinely processed, paraffin-embedded specimens. Following incubation with the primary rabbit antibody to human HER-2/neu protein, the kit employs a ready-to-use dextran-based visualization reagent. This reagent consists of both secondary goat anti-rabbit antibody molecules with horseradish peroxidase molecules linked to a common dextran polymer backbone, thus eliminating the need for sequential application of link antibody and peroxidase conjugated antibody. Enzymatic conversion of the subsequently added chromogen results in formation of visible reaction product at the antigen site. The specimen is then counterstained; a pathologist using light-microscopy interprets results.

FISH analysis at SHMC/PAML is performed using the FDA-approved PathVysion™ HER-2/neu DNA probe kit, produced by Vysis, Inc. Formalin fixed, paraffin-embedded breast tissue is processed using routine histological methods, and then slides are treated to allow hybridization of DNA probes to the nuclei present in the tissue section. The Pathvysion™ kit contains two direct-labeled DNA probes, one specific for the alphoid repetitive DNA (CEP 17, spectrum orange) present at the chromosome 17 centromere and the second for the HER-2/neu oncogene located at 17q11.2-12 (spectrum green). Enumeration of the probes allows a ratio of the number of copies of chromosome 17 to the number of copies of HER-2/neu to be obtained; this enables quantification of low versus high amplification levels, and allows an estimate of the percentage of cells with HER-2/neu gene amplification. The clinically relevant distinction is whether the gene amplification is due to increased gene copy number on the two chromosome 17 homologues normally present or an increase in the number of chromosome 17s in the cells. In the majority of cases, ratio equivalents less than 2.0 are indicative of a normal/negative result, ratios of 2.1 and over indicate that amplification is present and to what degree. Interpretation of this data will be performed and reported from the Vysis-certified Cytogenetics laboratory at SHMC.

## References


1. Wingo, P.A., Tong, T., Bolden, S., "Cancer Statistics", 1995;45:1:8-31.
2. "Cancer Rates and Risks", 4th ed., National Institutes of Health, National Cancer Institute, 1996, p. 120.
3. Slamon, D.J., Clark, G.M., Song, S.G., Levin, W.J., Ullrich, A., McGuire, W.L. "Human breast Cancer: Correlation of relapse and survival with amplification of the her-2/neu oncogene". Science, 235:177-182, 1987.
4. Xing, W.R., Gilchrist, K.W., Harris, C.P., Samson, W., Meisner, L.F. "FISH detection of HER-s/neu oncogene amplification in early onset breast cancer". Breast Cancer Res. And Treatment 39(2):203-212, 1996.
5. Press, M.F. Bernstein, L., Thomas, P.A., Meisner, L.F., Zhou, J.Y., Ma, Y., Hung, G., Robinson, R.A., Harris, C., El-Naggar, A., Slamon, D.J., Phillips, R.N., Ross, J.S., Wolman, S.R., Flom, K.J., "Her-2/neu gene amplification characterized by fluorescence in situ hybridization: poor prognosis in node-negative breast carcinomas", J. Clinical Oncology 15(8):2894-2904, 1997.


*Provided for the clients of*

PATHOLOGY ASSOCIATES MEDICAL LABORATORIES
PACLAB NETWORK LABORATORIES
TRI-CITIES LABORATORY
TREASURE VALLEY LABORATORY

*For more information, please contact your local representative.*